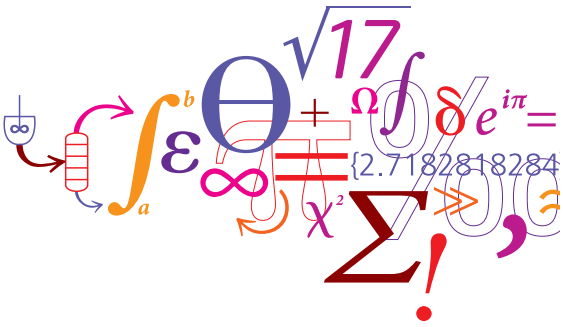
Computer-aided modeling for efficient and innovative product-process engineering

Heitzig, Martina

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

Computer-aided modelling for efficient and innovative product-process engineering



Martina Heitzig
Ph.D. Thesis
December 2011

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Computer-aided modelling for efficient and innovative product-process engineering

Ph. D. Thesis

Martina Heitzig

December 2011

Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark

DTU Chemical Engineering
Department of Chemical and Biochemical Engineering

Address: Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark
Building 229
DK-2800 Kgs. Lyngby
Denmark
Phone: +45 4525 2800
Fax: +45 4588 4588
Web: www.capec.kt.dtu.dk

Print: J&R Frydenberg A/S
København
june 2012

ISBN: 978-87-92481-73-3

Preface

This thesis is submitted as partial fulfilment of the requirements for the Doctor of Philosophy (*Ph.D.*) *degree at the Technical University of Denmark (DTU). The work presented has been carried out at the Computer Aided Process-Product Engineering Center (CAPEC) at the Department of Chemical & Biochemical Engineering from November 2008 till December 2011 under the guidance of Professor Rafiqul Gani as main supervisor as well as Associate Professor Gürkan Sin and Professor Peter Glarborg as co-supervisors. The project was funded by a scholarship from the Technical University of Denmark.*

I have the pleasure to acknowledge numerous persons who have contributed directly and indirectly to the development of this project:

I would like to start by expressing my special gratitude to my supervisors Rafiqul Gani, Gürkan Sin and Peter Glarborg for their valuable support, inspiring ideas and enthusiasm for my project and for providing me with so many opportunities to broaden my mind and meet with interesting people related to my field of study.

I also wish to thank my collaborators throughout the project. Special thanks to Professor Mauricio Sales-Cruz form the Universidad Autónoma Metropolitana (UAM), Mexico City for helping me getting started with the programming work on the software and supporting my external research stay in his group. Furthermore, I am grateful to Dr. Chris Gregson and Yunhong *Rong from Firmenich Inc. for providing me with an interesting case study dealing with fragrance aerosol systems and performing the required experiments for the validation of the developed models. Many thanks to Professor Andreas Linninger and Cierra Hall from the University of Illinois at Chicago (UIC) for the collaboration and fruitful discussions during the pharmacokinetic case study and for hosting me to work with them at UIC.*

A very warm thank you to all my co-workers and friends at CAPEC and PROCESS who have made these last three years a very special and pleasant time. I have always enjoyed the friendly atmosphere, discussions, lunch times and the many social events.

I am also very grateful to my parents Christa and Wilhelm and brothers Stefan, Matthias and Fabian for always being there and supporting me.

Finally and especially, I would like to thank my boyfriend Philip for his love, the nice conversations, great times and huge amount of patience!

Kgs. Lyngby, December 2011

Martina Heitzig

Abstract

Model-based computer aided product-process engineering has attained increased importance in a number of industries, including pharmaceuticals, petrochemicals, fine chemicals, polymers, biotechnology, food, energy and water. This trend is set to continue due to the substantial benefits computer-aided methods provide. The key prerequisite of computer-aided product-process engineering is however the availability of models of different types, forms and application modes. The development of the models required for the systems under investigation tends to be a challenging, time-consuming and therefore cost-intensive task involving numerous steps, expert skills and different modelling tools. The objective of this project is to systematize the process of model development and application thereby increasing the efficiency of the modeller as well as model quality.

The main contributions of this thesis are a generic methodology for the process of model development and application, combining in-depth algorithmic work-flows for the different modelling tasks involved and the development of a computer-aided modelling framework. This framework is structured, is based on the generic modelling methodology, partially automates the involved work-flows by integrating the required tools and, supports and guides the user through the different work-flow steps. Supported modelling tasks are the establishment of the modelling objective, the collection of the required system information, model construction including numerical analysis, derivation of solution strategy and connection to appropriate solvers, model identification/ discrimination as well as model application for simulation and optimization. The computer-aided modelling framework has been implemented into an user-friendly software.

A variety of case studies from different areas in chemical and biochemical engineering have been solved to illustrate the application of the generic modelling methodology, the computer-aided modelling framework and the developed software tool.

Resumé på Dansk

Model baserede computer understøttet produkt process engineering har opnået øget betydning i forskelligste industrielle brancher som for eksampel farmaceutisk produktion, petrokemi, finkemikalier, polymerer, bioteknologi, fødevarer, energi og vand. Denne trend er forventet at fortsætte på grund af substantielle fordele, hvilke computer understøttede metoder medfører. Den primære forudsætning af computer understøttet produkt process engineering er selvfølgelig den tilgængelighed af modeller af forskellige typer, former og anvendelser. Udviklingen af den påkrævet modellen for de undersøgte systemer er normalt en tidskrævende udfordring og derfor mest også dyrt. Den involverer forskelligste trin, fagekspert viden og dygtighed og forskellige modellerings værktøjer. Formålet af dette projekt er at systematisere den model udviklings proces og anvendelse og dermed øge effektiviteten af modeller såvel som kvaliteten. Den væsentlige bidrag af denne PhD afhandling er en generisk metodologi for proces model udviklingen og anvendelse i kombination med grundige algoritmiske arbejdes diagrammer for de forskellige involverede modeller opgaver og udviklingen af computer understøttede modeller rammer hvilke er strukturbaseret på den generiske metodologi, delvis automatiseret i de forskellige arbejdstrin og kombinerer alle påkrævet værktøjer, understøttelse og vejledning for de forskellige arbejdstrin. Understøttede modelleringsopgaver er etableringen af modeller mål, indsamling af de nødvendige informationer, model formulering inklusive numeriske analyser, etablering af løsningsstrategier og forbinding med den passende løsningsmodul, model identificering og sondering såvel som model anvendelse for simulation og optimering. Den computer understøttede modeller ramme blev implementeret i en brugervenlig software. En række forskellige demonstrationseksempler fra forskellige områder i kemisk og biokemiske engineering blev løst for udvikling og validering af den generiske modellerings metodologi og den computer understøttet modeller ramme anvendt på den udviklet software værktøj.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

The main objective of this PhD-project is the development of a structured computer-aided modelling framework that is based on a modelling methodology, which combines in-depth work-flows and data-flows for different modelling tasks related to model development and application with the goal to systematize the modelling process and to increase the efficiency of the modeller. For each work-flow step the support and features that a computer-aided modelling framework can provide as well as the potential for automation of the work-flow needs to be identified. Finally, the developed computer-aided modelling framework needs to be implemented into an user-friendly software.

In order to highlight the motivation for this PhD-project this chapter starts with a general discussion about the importance and challenges of computer-aided modelling (Section 1.1) and continues with a review on modelling methodologies and modelling tools (Section 1.2). Based on this review the identified issues and needs for a systematic computer-aided modelling methodology and its corresponding modelling tools are discussed (Section 1.3). These issues and needs give a brief overview on the shape of the computer-aided modelling framework to be developed. Finally, the objectives of the Ph.D. project (Section 1.4) and the structure of the thesis (Section 1.5) are summarized.

First, however, the definitions of important terms used throughout this thesis are given:

- *Model:* 'The representation of a real or virtual physico-chemical, economic, social or human situation, in an alternate mathematical or physical form, for an envisaged purpose' (Cameron & Gani, 2011).
- *Modelling methodology*: Methodology for the process of model development and application , that is, representation of the modelling process in terms of a ordered set of tasks and sub-tasks.
  In this work, a methodology based on work-flows and data-flows for the different sub-tasks of the modelling process is proposed.
- *Work-flow*: A work-flow summarizes the different steps required to complete a given task. In this context 'in-depth' provides a detailed explanation of each work-flow step and the corresponding sub-steps in an algorithmic form.

- *Data-flow*: Required data/information for different work-flow (sub-)steps as well as output data/information.
- *Computer-aided modelling framework*: A computer-aided modelling framework provides the architecture through which the computer-aided methods and tools can be implemented and used according to the work-flow and data-flow of the methodology.
- *Modelling tool/software*: Actual implementation of computer-aided modelling framework in a software.

1.1 Computer-aided modelling: importance and challenges

Computer-aided modelling and the related methods and tools play a role of increasing importance in product-process design across a number of industries including pharmaceuticals, petrochemicals, fine chemicals, polymers, biotechnology, food, energy and water. This is for a very simple reason: Model-based computer-aided methods are of great value in pushing forward the development of the improved products and/or the new innovative processes in order to master the enormous current and future challenges related to feedstock shortages, increasing population growth, environmental issues, safety regulations, demand for increasing product quality, sharper competition due to globalization and shorter product-lifetimes.

The main contributions of computer-aided model-based methods and tools are the prediction and optimization of product-process behaviour, their potential to replace resource-demanding, time-consuming and cost-intensive experiments, and, to deliver truly innovative solutions that are not necessarily obtained by conventional approaches (experimental and/or computational trial and error investigations). Another very important benefit is a better understanding and analysis of the domain system. For example, in product design, model-based systems can be used to either find promising candidates for final verification by experiments, or, to evaluate and analyse existing products to identify opportunities for improvement. To this end, the concept of a virtual lab for product-process design (Morales-Rodríguez & Gani, 2007) is interesting as it allows the product designer to perform virtual experiments to find, select and/or evaluate products and the important processing steps that define their production. The importance of model-based computer-aided techniques has also been emphasized by a number of authors. For example, Grossmann & Westerberg (2000) as well as Pantelides (2001) identified modelling, simulation, and optimization of large-scale systems to be crucial for handling the complexity of products and their related chemical processes.

However, the development, analysis and identification of the required reliable, validated and predictive models and their efficient application are the key prerequisites for a model-based

approach. As these models may be complex and often require multiple time and/or lengths scales, their development and application for product-process design is a non-trivial, time-consuming, and therefore, cost-intensive task involving numerous steps, expert skills (process knowledge, numerical analysis, statistical methods, etc.) and different tools. In fact, Foss et al. (1998) state that the effort spent for modelling is the most time-consuming factor in an industrial project involving model-based process engineering techniques, even though there are a large variety of commercially available modelling tools. This is confirmed by Preisig (2010) who states the development of a plant model to be one of the major bottle necks in engineering. Foss et al. (1998) further warn that this prevents the application of state-of-the-art model based technology in industrial projects. Likewise, Klatt et al. (2009) emphasize the need to combine and implement the developed modelling methods in user-friendly modelling tools.

The above observations have motivated this PhD-project to develop a generic computer-aided modelling methodology, which systematizes the modelling process, supports the modeller in the development and application of the required models and incorporates state-of-the-art modelling techniques. Such a modelling methodology aims to reduce the time and resources needed for the development of consistent models and their application, thereby reducing the overall time for product process development.

1.2 Overview of computer-aided modelling methodologies and tools

The state-of-the-art of modelling methodologies (Section 1.2.1) and existing modelling tools (Section 1.2.2) is reviewed in order to identify issues and needs for further development.

1.2.1 Modelling methodologies

In the last decade, the modelling process has been studied by a number of authors since a detailed understanding of this lays the foundation for the development of modelling tools that are effective in overcoming the problems and challenges related to model development and application (Foss et al., 1998). This has led to the development of modelling methodologies that divide the modelling process into a collection of tasks.

1.2.1.1 Modelling process

For a study of an industrial process, Foss et al. (1998) have identified eight tasks of the modelling process. Also Cameron & Gani (2011) have divided the modelling process into eight tasks. Both approaches are shown in Figure 1.1.

Figure 1.1 *Tasks of modelling process; left: Foss et al. (1998), right: Cameron & Gani (2011).*

Although in both cases, the reported modelling process is very similar in its main tasks and communicated sub-tasks, the order of the tasks differs (e.g. experimental data). The first task identified in the modelling process is the problem definition. Foss et al. (1998) continue with the selection of the environment for model implementation and afterwards go to the development of the conceptual model whereas Cameron & Gani (2011) do not consider model environment selection at this point of the modelling process. In general, a conceptual model does not depend on the software environment applied for later implementation, analysis and solution of the

model. For this reason, it is not required to select the modelling environment before developing a conceptual model. However, in some cases the software environment may offer support during the conceptual model derivation. After the conceptual modelling task, Cameron & Gani (2011) suggest to proceed with the data requirement task that has already been considered by Foss et al. (1998) during the problem statement task. For both methods, model construction is the next task. At this point the modelling process proposed by Cameron & Gani (2011) demands the selection of a modelling environment. Foss et al. (1998) merge the model construction task with the numerical model analysis and verification tasks while Cameron & Gani (2011) consider separate tasks for model solution and verification. The next task is the model validation where the performance of the model with respect to the modelling goal is evaluated. Cameron & Gani (2011) consider parameter estimation as a possible sub-task of model validation while Foss et al. (1998) do not consider parameter estimation. In general, parameter estimation is an important part of the modelling process and should not be omitted. The modelling process is iterative and the steps given in Figure 1.1 need to be repeated until the model performance achieves the modelling goal.

Preisig (2010) proposes a 13-step methodology specifically for the construction of models for physical, chemical or biological systems which is summarized in Table 1.1.

Table 1.1 *Process model construction sequence proposed by Preisig (2010).*

Operations of model construction for process models	***More details:***
0. Sketch process.	Drawing of system and boundaries
1. Map process into an abstract network of communicating control volumes.	Nodes of network represent capacities (smallest control volumes), arcs represent connection flows
2. Add type information to nodes and arcs	physical system, types of flows (mass, heat, work), etc.
3. Establish base model by establishing the balances of the relevant conservation equations for each node.	
4. Add description of the transfer defined in the network	e.g. conductive heat transfer, radiation, diffusional or convective mass flow
5. Add internal dynamics for each node	Reaction rate and phase transition models
6. Obtain the secondary state variables introduced by the transfer and transposition kinetics	
7. Add control	
8. Make appropriate simplifying assumptions	e.g. constant properties, constant volume or pressure, very fast transport or reaction
9. Instantiation of mathematical problem	Instantiate parameters and conditions; variable values,
10. Generate code for solver	
11. Solving	
12. Verifying and identifying model	by comparing plant with simulated behaviour

Compared to the methodology proposed by Cameron & Gani (2011) and Foss et al. (1998) Preisig (2010) does not consider the establishment of a modelling goal and collection of system information.

Morbach et al. (2009) have proposed the formal ontology 'OntoCAPE' for the domain of computer-aided process engineering with the intention to support 'the construction of software tools for different tasks such as knowledge management, mathematical modelling or plant design'. The authors state that a domain ontology is a framework that represents the knowledge of an entire application domain. OntoCAPE is structured according to 5 layers with different levels of abstraction: Meta layer, upper layer, conceptual layer, application-oriented layer and application-specific layer (see Table 1.2 for more detailed description). The upper layers provide more general knowledge about ontologies and systems theory while the knowledge provided by the lower layers is domain-related.

Table 1.2 *5 Layers of OntoCAPE – ontology for computer-aided process engineering (Morbach et al., 2009).*

Layer	***Description***
1. Meta layer	▪ Fundamental modelling concepts ▪ Design principles of OntoCAPE ➔ 'meta layer supports ontology engineering and ensures a consistent modelling style by providing guidance for the extension and/or modification of the ontology' (Morbach et al., 2009)
2. Upper layer	▪ Principles of general systems theory and systems engineering ▪ Systems-theoretical and physicochemical primitives and their relations, e.g. System, Property, Value, PhysicalQuantitity, UnitOfMeasure, etc
3. Conceptual layer	▪ Establishment of conceptual model of CAPE domain (covers unit operations, equipment and machinery, materials and their thermo-physical properties, chemical process behaviour, modelling and simulation, etc.)
4. Application-oriented layer	▪ Classes and relations for practical application of ontology: ➔ General extension of ontology towards certain application areas; provides for example: chemical species data for atoms, molecules and polymers, description of typical process units, property models, establishment of customary mathematical models for process units (e.g. ideal reactor models, tray-by-tray models for distillation columns)
5. Application- specific layer	▪ Classes and relations for practical application of ontology: ➔ specialized classes and relations for concrete applications

The layers contain different modules which 'assemble a number of interrelated classes, relations, and axioms, which jointly conceptualize a particular topic´. Related modules are summarized in 'partial models' which can extend over several layers. With respect to computer-

aided modelling the partial model 'model' which 'defines notions required for a description of mathematical models and model building' is of special interest. It contains the following modules: numerical_solution_strategy, equation_system, mathematical_model, cost_model, process_model, property_models, laws and process_unit_model. OntoCAPE has an informal presentation which consists of technical reports as well as a formal representation based on the Ontology Web Language (OWL) (Bechhofer et al., 2004). OntoCAPE has been applied to develop different software tools, for example for the automated creation of process models based on a conceptual model provided by the user (Yang & Marquardt, 2004).

1.2.1.2 Multi-scale modelling

Multi-scale modelling has gained increasing importance due to the fact that it has a large potential to improve the ability of modelling complex systems that include several length and/or time scales with higher detail and accuracy as well as to extend the application areas of models developed by chemical engineers (Ingram et al., 2004).

Ingram et al. (2004) have defined a multi-scale model as 'a composite mathematical model formed from two or more sub-models that describe phenomena at different scales.' In this definition the term sub-model relates to a 'component model that describes only one scale of the system'. A scale in turn is given by 'some characteristic time or length of the objects and phenomena involved'. Furthermore, a model that describes a system on different levels of detail can also be characterized as a multi-scale model.

Yang & Marquardt (2009) define a scale or a level as a 'part of a certain decomposition scheme'. The authors further state that 'in a multilevel system, one level is said to be "lower" than another one, if any component in the former is always part of a particular component in the latter. A system is defined as a 'set of things and their connections'. A thing is 'the pair of a substantial individual and the set of its properties' with a substantial individual being an 'individual that exists'. Yang & Marquardt (2009) state a reactor to be an example of a thing which refers to the reactor as the individual and its properties (e.g. volume, yield). A component is defined as any element within the composition of a system while the composition of a system is the set of things in the system. Based on these concept Yang & Marquardt (2009) define a single-scale model to be 'a collection of the laws of all components at a certain level of observation and the laws of coupling among these components' and a multi-scale model as 'a collection of single-scale models as well as laws characterizing inter-scale relations of the system being modeled'.

Németh et al. (2005) state that the different size-scales of relevance for chemical engineering applications span a range from the quantum mechanical level (10^{-13} m) with time-scales of 10^{-16} seconds to a global scale ($\geq 10^{4}$ m, time-scale $\geq 10^{8}$ s). Figure 1.2 gives a more detailed

overview of different size and time scales relevant for modelling in chemical and biochemical engineering, starting from the molecular level (Grossmann, 2004).

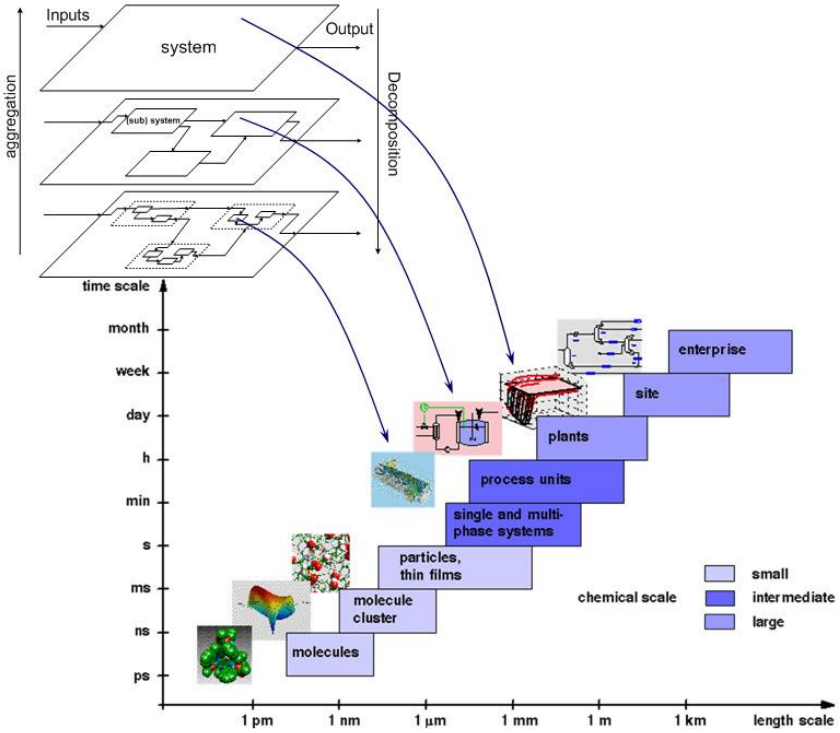


Figure 1.2 *Size scale of multi-scale models in chemical engineering (from Grossmann, 2004).*

In order to deal with the additional complexity multi-scale modelling introduces to the modelling process, the above described modelling methodologies need to be extended. To this end, Ingram et al. (2004) have identified three tasks that are specific to multi-scale modelling:

1. Identification of the time and length scales of importance for the multi-scale model;
2. Derivation of the sub-models for each scale (If the models do not exist they need to be developed applying a systematic modelling methodology like given in Figure 1.1.);
3. Linking of the sub-models to form the final multi-scale model.

The scales to be included in the multi-scale scenario are determined by the following factors (Ingram et al., 2004):

- Scales associated to the required input and output variables;
- Unknown/ uncertain variable values in existing scales;
- Unsatisfactory model predictions;
- Experience of modeller (aided by domain-specific scale maps);
- Geometry of system;

In case the performance of the final linked multi-scale model is not satisfactory Ingram (2005) proposes the following options for improvement:

- Add new scales;
- Remove scales (in case the modelling objective is still met and the model implementation is too large or the solution too slow);
- Modify an existing scale by dividing it into sub-scales.

In order to construct a multi-scale scenario Ingram et al. (2004) have proposed different strategies, which can be selected according to the specific modelling problem:

1. Bottom-up: Starting from the smallest scale of interest sub-models of increasing scales are gradually added to the model.
2. Top-down: Starting from the highest scale of interest the degree of detail of the model is increased by gradually adding lower scale models.
3. Concurrent: All scales of interest are added simultaneously.
4. Middle-out: Starting with a selected scale (e.g. because most detailed knowledge available for this scale) the model is extended by gradually adding smaller and bigger scales.

Finally, Ingram et al. (2004) propose a classification of multi-scale models based on five alternative integration frameworks of two scales, which is summarized in Table 1.3. The classification is an extension of the work of Pantelides (2001). The frameworks contribute to systematically categorizing the different types of linking schemes for multi-scale models and elucidating the information transfer between the scales.

Table 1.3 *Classification of multi-scale models based on five different integration frameworks (valid for 2 scales), adapted from Ingram et al., 2004.*

Integration framework	explanation	example
Multidomain	'Adjacent, largely non-overlapping micro- and macro-scale[1]-simulated regions that communicate across interface'	macro-scale[1]: bulk model, micro-scale[1]: pellet model
Embedded	'Micro-scale model is formally embedded in macro-scale model.'	macro-scale: average values from bulk and pellet, micro-scale: pellet model
Parallel	'Both models span the system domain. They may both have macro-scale and micro-scale features, however the models are complementary in the detail with which they represent the controlling phenomena'	Scale 1: detailed hydrodynamic model with simple kinetics, Scale 2: detailed kinetic model with simple flow pattern
Serial:		
(i) **Simplification**	'the order of the micro-scale model is reduced then it is used by the macro scale model'	
(ii) **Transformation**	'the micro-scale model is formally transformed into a macro-scale model'	
(iii) **One-way-coupling**	'the nature of the system is such that there is strictly one-way flow of information from the micro-scale to the macro-scale, consequently the micro-scale model is solved then its results are used in the macro scale model'	aerosol case study (see Section 5.4)
Simultaneous	'The entire system is solved at the micro scale'	molecular simulations

[1] *For this table the terms 'macro' and 'micro' do not relate to the actual size of the scales but to the relative size of the scales. That means that 'macro' stands for the bigger scale and 'micro' for the smaller scale.*

Yang & Marquardt (2009) establish a distinct categorization of multi-scale models in three different types based on inter-scale laws: 1. Scale-collecting models, 2. Scale-connecting models and 3. Scale-integrating models. First, the inter-scale laws defined by Yang & Marquardt (2009) are introduced. Second, the three different multi-scale model types identified by Yang & Marquardt (2009) based on the inter-scale laws are defined.

1. Aggregation law: 'Let σ be a system with levels $V = (V_k | 1 \leq k \leq n)$. Let x be a component on level $V_i, i > 1$ with an immediate composition $C^{im}(x)$. An aggregation law on x, denoted $l^a(x)$, is a relation that maps one or more state functions of all the components in $C^{im}(x)$ to a state function of x'. The level i=1 is always the lowest level.

The state functions give the properties of a thing. The immediate composition $C^{im}(x)$ of a component x on level $V_i, i > 1$ is defined as 'a set of components on levels preceding V_i.such that, each component in this set does not belong to the composition of any component located in a level below V_i'.

2. Disaggregation law: 'Let σ be a system with levels $V = (V_k | 1 \leq k \leq n)$. Let x be a component on level $V_i, i > 1$ with an immediate composition $C^{im}(x)$. A disaggregation law on x, denoted as $l^d(x)$, is a relation that maps a state function of x to a state function of all the components in $C^{im}(x)$'.
3. Mereological connection law: 'Let σ be a system with levels $V = (V_k | 1 \leq k \leq n)$. Let x be a component on level $V_i, i > 1$ with an immediate composition $C^{im}(x)$. A mereological connection law on x, denoted by $l^m(x)$, is a relation between a coupling-induced state function of x […] and one or more coupling-induced state functions of a set of components $U \subseteq C^{im}(x)$'. A coupling-induced state function is any state function of x iff it represents a property of x that occurs only due to the coupling of x with another thing y.

Having explained the different interscale laws the three types of multi-scale models proposed by Yang & Marquardt (2009) are elaborated on, as shown below.

Type 1: A scale collecting model does not include 'neither an aggregation law, a disaggregation law, nor a mereological law on any level'.

Type 2: A scale-connecting model contains 'one or more mereological connection laws but no aggregation law or disaggregation law'.

Type 3: A scale-integrating includes 'at least one aggregation law or one disaggregation law'.

Yang & Marquardt (2009) further show that the multi-scale model types given by the different integration frameworks introduced by Pantelides (2001) and Ingram et al. (2004) are special cases of the three multi-scale model types given by the proposed categorization based on interscale laws.

1.2.2 Modelling tools

According to von Wedel et al. (2002) and Marquardt (1996) the existing modelling tools can be structured into three main groups: programming languages, generic modelling languages and domain oriented tools. In the following, a detailed explanation of the three groups including examples is given (this has been adapted from von Wedel et al. (2002)). The first group includes

the *programming languages* like Fortran, C, C++, etc. The second group is formed by the *generic modelling languages* that support the modeller in formulating the problem but do not provide any domain specific concepts. Wedel et al. (2002) classify generic modelling languages into two sub-groups: Mathematical modelling languages and system modelling languages. The former simplify the mathematical formulation of a problem but do not provide the means for structuring the resulting sets of equations. Examples are GAMS (Brooke et al., 1998) and MathML (Ausbrooks, et al., 2001). In contrast to the mathematical modelling languages, the systems modelling languages consider a model as a part of an overall system where model decomposition and aggregation are two main features. A representative of the systems modelling languages is Modelica (Modelica Association, 2000) a standardized modelling language based on object-oriented concepts. Other examples are gPROMS (Process Systems Enterprise, 2011) and Custom Modeler (Aspentech, 2011a). According to *domain-oriented tools* (third group of modelling tools), the model development process is based on providing concepts instead of equations and the tool generates the equations based on the user-specifications. Two different subgroups of domain-oriented tools can be distinguished. The first subgroup is formed by the so-called flowsheeting tools such as Aspen Plus (Aspentech, 2011b) and Pro II (Simsci, 2011), which provide libraries with models for different unit operations that the user can combine to build a process. These tools provide a maximum of domain-specific support and allow fast and efficient development of process models, connection to thermodynamic databases and process simulation without writing any equations. The drawback is the limited flexibility since the user relies to a great extent on the models available in the libraries and has not much insight in the model equations and solution process.

Providing more flexibility and allowing the consideration of phenomena on a higher degree of detail has motivated the development of the second subgroup of domain-oriented tools, the process modelling languages (von Wedel et al., 2002). They are based on the decomposition approach that decomposes the flowsheet not only in its unit operations but also on levels below the unit scale. Examples for this subgroup are MODEL.LA (Stephanopoulos et al., 1990) and ModDev (Jensen & Gani, 1996; Jensen, 1998). A more extended review on existing modelling tools has been given by Sales-Cruz (2006).

Von Wedel et al. (2002) propose a 3-layer approach for the contents and functionalities of a modern modelling tool consisting of a mathematical base layer, a systems engineering layer, and a chemical engineering layer to introduce the domain knowledge. In conclusion it can be stated that the systematic computer-aided modelling tool box combines elements of the three groups of modelling tools mentioned above and incorporates them through a framework that provides a balance between support, automation and flexibility (Heitzig et al., 2011a).

1.3 Issues and needs

1.3.1 Modelling methodologies

The existing modelling methodologies require extensions by increasing the degree of detail and adding in-depth work-flows and data-flows for the different tasks of the modelling process as, for example, identified by Foss et al. (1998) and Cameron & Gani (2011). Preisig (2010) states: 'Is it not surprising that there has been so little of a visible attempt to make modelling – […] – a structured, well-defined process?' It is also important to complete the overall picture of the modelling process by including in-depth work-flows and data-flows for tasks like model identification, discrimination, validation as well as application (simulation and optimization) and combining all work-flows to one overall modelling methodology. Furthermore, a work-flow for multi-scale model construction that considers the three specific multi-scale modelling tasks identified by Ingram et al. (2004) has not yet been developed. Klatt & Marquardt (2009) state that 'research on modeling methodologies should be of primary interest' to the PSE discipline and, in this context, name multi-scale modelling as one of the areas where more research is required. The modelling methodology resulting from the integration of the in-depth work-flows for the different modelling tasks will communicate the big picture of model development and provide more detail at the same time. It needs to be transformed into a computer-aided modelling framework and software tool.

1.3.2 Modelling tools

Improving and extending of existing modelling tools can have a number of benefits related to:

- Reduction of overall time required to develop a model and generate results (e.g. Preisig, 2010), by e.g. reduction of modelling errors and programming effort (Kuntsche et al., 2011).
- Reduction of modelling errors (Kuntsche et al., 2011)/ improvement of model quality.

Foss et al. (1998) emphasize that the detailed understanding of the model development process is the basis for the development of advanced modelling tools that truly enhance the 'productivity' of the modeller and the 'quality' of the models. Others agree with this statement (e.g. Sales-Cruz, 2006; Preisig, 2010). Klatt & Marquardt (2009) identify further potential for improving existing modelling tools. They see the lack of implementation of state-of-the-art modelling techniques as one of the major shortcomings.

A literature review on existing modelling tools as well as the needs for improving these tools has been performed. It has revealed 16 key areas where a modelling tool should provide support. These areas are summarized in Table 1.4

Table 1.4 *Key areas for a modelling tool to provide support.*

Key features of modelling tool	***references***
1 Structured development of models based on the modelling process, automation of part of the modelling process	Lohmann & Marquardt (1996); Bogusch et al. (2001); Sales-Cruz (2006); Preisig (2010)
2 Documentation of model and modelling process (including modelling objective, assumptions, system information)	Foss et al. (1998); Bogusch et al. (2001); Sales-Cruz (2006); Kuntsche et al. (2011)
3 Model re-use	Foss et al. (1998); Bogusch et al. (2001)
4 Model libraries	Foss et al. (1998); Bogusch et al. (2001); Sales-Cruz (2006); Process Systems Enterprise (2010a-d; 2011); Aspentech (2003; 2011a)
5 Model decomposition	Marquardt (1996); Foss et al. (1998); Hangos & Cameron (2001); Sales-Cruz (2006)
6 Model aggregation	Sales-Cruz (2006); Process Systems Enterprise (2010a-d; 2011); Aspentech (2003; 2011a)
7 Support for equation generation	Jensen & Gani (1996); Jensen (1998); Foss et al. (1998); Bogusch et al. (2001); Preisig (2010)
8 Simple implementation of model equations	Sales-Cruz (2006); Kuntsche et al. (2011)
9 Numerical model analysis	Foss et al. (1998); Sales-Cruz (2006); Aspentech (2003; 2011a)
10 Model verification/ debugging	Foss et al. (1998)
11 Systematic model reduction/ simplification	Foss et al. (1998); Preisig (2010);
12 Model identification/ validation	Foss et al. (1998); Sales-Cruz (2006); Process Systems Enterprise (2010a-d; 2011); Aspentech (2003; 2011a)
13 Simulation	Foss et al. (1998); Sales-Cruz (2006); Process Systems Enterprise (2010a-d; 2011); Aspentech (2003; 2011a)
14 Optimization	Foss et al. (1998); Process Systems Enterprise (2010a-d; 2011); Aspentech (2003; 2011a)
15 Support for multi-scale modelling	Ingram et al. (2004); Klatt & Marquardt (2009); Process Systems Enterprise (2010a-d; 2011)
16 Domain knowledge/support	Bogusch et al. (2001); Jensen & Gani (1996); Jensen (1998); Preisig (2010); Aspentech (2003, 2011a)

Based on the 16 key areas existing modelling tools have been investigated with respect to their main features in order to identify needs and potential for improvement. Table 1.5 gives an overview of the main features and gaps for model development and application of four different modelling tools, two of which come from academia. MOSAIC (Kuntsche et al., 2011) is the most recently developed modelling tool reported in literature. ProcessModeller (Preisig, 2010) has been chosen because it is based on the methodology for generation of an implemented model presented in Table 1.1. The two commercial modelling tools presented here are Aspen Custom Modeler (Aspentech, 2003; 2011a) and gPROMS (Process Systems Enterprise, 2010a-d; 2011), which are the most commonly used custom modelling tools for chemical engineering applications. Table 1.5 provides a short description for each tool, lists the main features/advantages as well as the main gaps/drawbacks. For each feature the key areas it addresses corresponding to Table 1.4 are given in brackets.

Table 1.5 *Main features and gaps of existing modelling tools.*

Tool and short description	***Main features/advantages*** *(the key areas are highlighted in parenthesis)*	***Main gaps/drawback***
MOSAIC (Kuntsche et al., 2011) ***Description:*** Web-based modelling tool close to documentation level with model equation input using Latex and automated export to selected target code (e.g. Matlab)	▪ Enter model equations using Latex in a syntax close to representation in scientific papers -> reduction of modelling errors ***(8, 10)*** ▪ Output of correct programming code for a selected and supported software environment (e.g. Matlab) -> useful for taking advantage of using optimal tool for each task, for multi-disciplinary collaboration (each partner uses own tools), minimize programming effort ***(8)*** ▪ Reusability feature of equation systems or equations from one model in another model (synonym lists to handle different notations in the models) ->reduction of modelling errors and programming effort ***(2, 3, 4, 5, 6, 7, 8)*** ▪ Each model/ equations system is connected to an obligatory notation for all present variables -> advantages for documentation and allows re-use of equations in other models with different notation ***(2, 3)*** ▪ Automatic documentation ***(2, 3)*** ▪ Model data bases are shared via internet -> coping with locally distributed work-places ***(3, 4)***	▪ No work-flow structure ▪ Latex syntax can be unhandy, not everyone familiar with it ▪ No combination of model trans-lation with partly automated numerical model analysis: ▪ No incidence matrix, optimal equation ordering, singularity check, etc. ▪ Only output is code in a selected target language/tool ▪ Tool does not provide own solvers, optimizers, model analysis features etc. ▪ Can only transfer models to different environment if model is coded in MOSAIC -> always extra modelling effort to do that. ▪ Does not support PDEs (and their discretization) ▪ Solvers and solver options need to be adjusted in target modelling tool after export ▪ Not useful stand-alone ▪ No features to automatically create simplified models (e.g. steady state, linearized) ▪ Modeller is forced to provide nomenclature

Table 1.5 *CON'T* (Main *areas for a modelling tool to provide support)*

Tool and short description	***Main features/advantages*** *(the key areas are highlighted in parenthesis)*	***Main gaps/drawback***
ProcessModeller Preisig (2010) ***Description:*** Modelling tool shaped based on an analysis of the process of generating a coded model (see Table 1.1)	▪ Shaped based on a work-flow for generation of a coded model (see Table 1.1) ***(1, 7, 8)*** ▪ Representation of model in form of a graph (nodes= balance volumes, =connection streams) **(7,8)** ▪ complexity is handled by enabling a hierarchical representation of the network **(7,8)** ▪ Assign types to flows and nodes, colour coding is generated according to type of flows (mass, heat, work, and also reactivity and diffusivity, steady state or dynamic, lumped or distributed) **(7,8)** ▪ Selection where to add which transfer models (e.g. conductive heat transfer, radiation, diffusion, convection) **(7,8)** ▪ Adding transposition models (reactions and phase transition models) **(7,8)** ▪ Tool creates model equations -> very efficient, less errors **(7,8)** ▪ Model simplification mechanisms available (e.g. order-of-magnitude assumptions of reaction rate or mass transfer rate) **(7,8, 11)** ▪ Parts of the model tree can be saved as sub-models in the library **(3,4,5,7)** ▪ Export of model to target code (supported e.g. Modelica, gProms) -> no coding errors ***(8)***	▪ Limited flexibility: 1) Modeller is limited to equations provided in the ontology, 2) exclusively for chemical engineering domain ▪ Other elements of modelling process (modelling objective, detailed documentation, identification, validation) and model application not possible; focus is on equation generation ->cannot be used stand-alone ▪ No analysis of incidence matrix ▪ No documentation interface, no automated report generation

Table 1.5 *CON'T* (Main *areas for a modelling tool to provide support)*

Tool and short description	***Main features/advantages*** *(the key areas are highlighted in parenthesis)*	***Main gaps/drawback***
Aspen Custom Modeler (Aspentech, 2003; 2011a) ***Description:*** Modelling tool which allows custom model development and easy integration of developed model within a flowsheet.	▪ Model equations are provided in an equation editor with defined syntax (during compilation: syntax check for errors), sub-models possible, intrinsic sub-functions e.g. for bubble point calculation, PDEs and their discretization is supported ***(8, 15)*** ▪ Numerical model analysis: degree of freedom and propositions of variables to specify ***(9, 10)*** ▪ Structural singularity check can be performed (but not automated and not in work-flow) ***(9, 10)*** ▪ Model libraries (for unit operation and stream types, etc.) and creation of libraries ***(3, 4)*** ▪ Typical flow-sheeting features: ->create flow-sheet by drag and drop based on library, component selection, thermo-dynamic databases, change units of measure globally, etc. ***(3, 6, 7,16)*** ▪ Solvers: dynamic and algebraic, creates tables and plots of simulation results, message window with diagnostic output from solvers during solution process. ***(9, 10, 13)*** ▪ Task list for simulation (e.g. flow rate changes during dynamic simulation) created with specific syntax (syntax checking during compilation) ***(13)*** ▪ Switch to steady state without changing model implementation ***(7, 13)*** ▪ Flow-sheet optimization with constraints ***(14)*** ▪ Parameter estimation using steady state and dynamic data, output of results and statistics ***(12)***	▪ No work-flows ▪ Syntax not much simpler than programming in Fortran (e.g. variable declarations needed) ▪ No interface to link submodels ▪ Syntax for PDEs and their discretization complex, discretized form of model equations is not visible to modeller ▪ No equation generation tool ▪ No incidence matrix analysis and optimized equation ordering (at least not explicitly for modeller) ▪ No documentation interface, no automated report generation (but creates output of simulation and other information like e.g. degree of freedom in tables and plots) ▪ No sensitivity analysis ▪ No identifiability analysis ▪ No uncertainty analysis ▪ No equation-by-equation de-bugging

Table 1.5 *CON'T* (Main *areas for a modelling tool to provide support)*

Tool and short description	***Main features/advantages*** *(the key areas are highlighted in parenthesis)*	***Main gaps/drawback***
gPROMS (Process Systems Enterprise, 2010a-d; 2011) ***Description:*** Modelling tool which allows custom model development and easy integration of developed model within a flowsheet.	▪ Model equations are provided in an equation editor with defined language, sub-models possible (can be constructed graphically or with gPROMS language), PDE discretization based on user specifications ***(8, 15)*** ▪ Equation ordering ***(8, 9)*** ▪ Stream types for connecting models in flowsheets ***(3, 6, 7)*** ▪ Topology view for graphical construction of flowsheet models (drag and drop existing component models and equate their model ports) ***(3, 6, 7)*** ▪ Model analysis at the start of each simulation to check if the model is well-posed and whether alternative specifications are required for the degrees-of-freedom. Results of that analysis are shown in generated solution report ***(9, 2, 13)*** ▪ Tasks (reusable part of operating procedure) can be defined (syntax similar to model syntax) ***(13)*** ▪ Schedules for simulations ***(13)*** ▪ Switch to steady state simulation without changing model implementation ***(7, 13)*** ▪ For each model execution a case is generated which safes the main information with respect to input, and generated results ***(2, 13)*** ▪ Dynamic optimization, mixed integer optimization ***(14)*** ▪ Parameter estimation, algebraic and dynamic data, statistical analysis (e.g. 2-D confidence ellipsoids), reports with all results ***(2, 12)*** ▪ Design of experiments ***(12)*** ▪ Interface to connect to external software (e.g. physical properties packages, CFD tools) ***(12, 13, 16)*** ▪ Uncertainty analysis ***(12, 13)***	▪ No work-flows ▪ Syntax for model implementation not much simpler than programming in Fortran (e.g. variable declarations needed, values and initial conditions are provided in editor with model equations) ▪ PDE and their discretizations demands for very detailed declarations in txt-syntax ▪ No equation generation tool ▪ Variable classification to satisfy degree of freedom needs to be done by hand in txt-syntax without inline DOF-calculation and singularity check ▪ No incidence matrix (at least not explicitly for modeller) ▪ A number of features need to be set-up in txt-syntax like the model itself: Solvers and solver options, outputs, optimization, parameter estimation (inter-face for parts of the set-up, e.g. experimental data interface), uncertainty analysis ▪ No sensitivity analysis ▪ No identifiability analysis ▪ Only standard intrinsic functions, e.g. cos, but not bubble point calculation, etc. ▪ No physical property interface

Analysing the information given in Table 1.5 reveals that, even after combining all the features of the four modelling tools the 16 key areas not covered. However, different tools support different areas to differing extents. The main categories where all modelling tools show issues/gaps are summarized by:

1) Structured model development and application, automation of part of the modelling process (key area 1, Table 1.4)
2) Documentation of model and modelling process (key area 2)
3) Multi-scale modelling (key area 15)
4) Simple implementation of model equations (introducing of model equations, translation and connection to solvers) (key area 8)
5) Numerical model analysis (key area 9)

The developed computer-aided modelling tool needs to address all five categories identified above.

1.4 Objectives and overview of developed modelling methodology and corresponding computer-aided modelling framework

Section 1.3 has identified five different categories where existing modelling tools show shortcomings in order to tackle the challenges related to the development of complex models required by industry and academia. The first category is related to the need for a computer-aided modelling framework that structures and systematizes the process of model development and application based on a profound modelling methodology and in addition, identifies potential for atomizing parts of the modelling process. To ensure this structure is the main focus of the Ph.D.-project. The core of such a computer-aided modelling framework is the establishment of a generic modelling methodology that considers all tasks related to the modelling process. Section 1.3.1 has emphasized the related issues and needs for modelling methodologies. The proposed methodology combines in-depth work-flows and data-flows for the different tasks involved in model development and application. These are:

1. Modelling objective and system information;
2. Single-scale model construction or multi-scale model construction;
3. Model identification/ discrimination;
4. Model application for simulation;
5. Model evaluation and validation;
6. Model application for optimization.

The computer-aided modelling framework is developed based on the established generic modelling methodology by identifying the required features, data-base and library connections for each step of the different work-flows. In addition, it is investigated how the computer can provide maximum guidance and insights in the theoretical backgrounds of the methods in each work-flow step and which parts of the modelling process can be (fully or partly) automated without loss of flexibility. In doing this, the other four categories identified in Section 1.3 for the potential of improving existing modelling tools, are automatically addressed. Figure 1.3 summarizes the target-benefits and architecture of the developed computer-aided modelling framework.

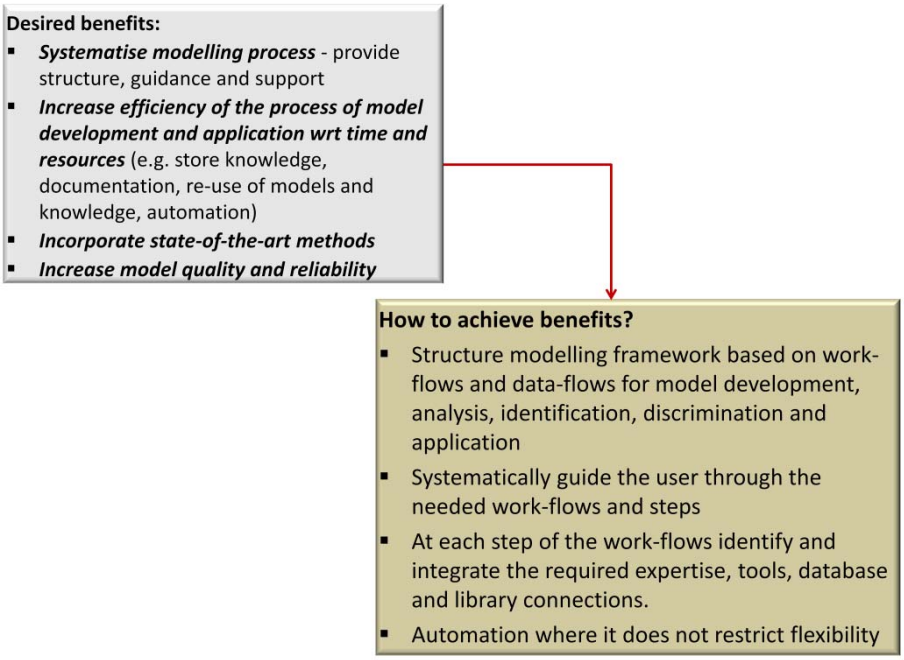


Figure 1.3 *Benefits and structure of developed computer-aided modelling framework.*

The computer-aided modelling framework has been developed in three steps. Firstly, the required work-flows and data-flows for the modelling methodology are identified and developed (see Chapter 2). This has been done based on the identified issues and needs and the gaps in the current state-of-the-art in modelling tools. A number of case studies (see Chapter 5) from different areas in chemical and biochemical engineering have been solved to validate and highlight the application of the methodology. Secondly, the required support, methods and tools that the computer-aided modelling framework needs to integrate for the established work-flows as well as the automation potential are identified and developed (see Chapter 3). Finally, the computer-aided modelling framework is implemented into an user-friendly software

(see Chapter 4). As proposed by von Wedel et al. (2002) the developed computer-aided modelling framework and modelling tool combine elements from all three groups of modelling tools (see 3-layer approach, Section 1.2.2).

1.5 Main objectives of PhD thesis

- ***Development of overall modelling methodology based on in-depth work-flows and data-flows for the different generic modelling tasks required for model development, analysis, identification, discrimination, documentation and application for simulation and optimization (see Chapter 2)*;**
    1. Identification of modelling tasks and their interconnection (based on literature, existing methodologies for the modelling process, case studies)
    2. Development of in-depth work-flows and data-flows for the modelling tasks including required methods (based on literature, existing work-flows and methods)
- ***Development of computer-aided modelling framework that is structured based on the modelling methodology by elaborating how the computer can support the modeller and making identified work-flows and data-flows computer-aided (see Chapter 3)* ;**
    1. Identification of required tools and features for each step in order to provide maximum support to the modeller
    2. Provide guidance and insights to theoretical background and application of the different methods and tools
    3. Analysis of opportunities for automation of steps without loss of flexibility
- **Development of software architecture and implementation of computer-aided modelling framework into user-friendly software (see Chapter 4);**
- **Validation of modelling methodology and computer-aided modelling framework based on the solution of case studies from very different areas in chemical and biochemical engineering (see Chapter 5).**

1.5 Structure of PhD thesis

This chapter of the thesis (Chapter 1) elaborates on the importance and challenges in computer-aided modelling and the state-of-the-art in the development of modelling methodologies and tools. Based on this the issues and needs for the extension of existing modelling methodologies and tools are identified. Finally, the main goals of the thesis are briefly summarized. Chapter 2 is dedicated to the description of the developed generic work-flow based modelling methodology

and elucidates the different work-flows and data-flows in detail. Chapter 3 presents the development of a computer-aided modelling framework based on the modelling methodology described in Chapter 2. The different computer-aided work-flows are presented highlighting the identified features a modelling tool should provide for each step of the work-flow together with the automation opportunities for different work-flow (sub-) steps. The goal of Chapter 4 is to present the developed modelling software based on the architecture of the modelling framework (see Chapter 3). Chapter 5 highlights the application of the modelling methodology and corresponding computer-aided modelling framework (and software) by solving several case studies from different areas in chemical and biochemical engineering. Finally, Chapter 6 summarizes the main achievements and conclusions of this Ph.D.-project and highlights important future challenges.

Chapter 2. Modelling methodology

This chapter presents the developed generic modelling methodology, which is based on the concept of decomposing the modelling process into a sequence of four phases for model development and one phase for model application. The main objective of the methodology is to systematically and efficiently generate a translated, analysed, identified, validated and reliable model with respect to a provided modelling objective and to apply the obtained model in various model-based studies of the modelled system. The model development process is an iterative procedure where different candidate models are developed, compared, evaluated and re-fined until the optimal model with respect to the modelling objective is found. The modelling objective might need extension and revision during model development and application.

A modelling problem involves the following tasks: the derivation of the model equations, their translation, analysis and finally, their solution. The generic model structure for chemical and biochemical engineering applications consists of (Cameron & Gani, 2011):

1. Balance equations/ conservation equations for mass, energy and momentum;
2. Constitutive equations/ models;
3. Conditional equations.

The balance equations are established for each balance volume in the system. For many applications not all three types (mass, energy, momentum) of balance equations are required. Conservation equations can be dynamic or steady state. Furthermore, they can be lumped or distributed. The need for constitutive equations/ models depends on the different phenomena (e.g. reaction) considered in the model. In many cases the constitutive equations are algebraic. They can also be dynamic equations, for example in modelling of diffusion related properties. The conditional equations refer to connection equations between different balance volumes of the system or with the surrounding. Another example for conditional equations is a closure constraint for mole fractions within a balance volume. Not every modelling problem needs to have all three equation types. In some cases a model consists only of constitutive equations and does not contain any balance equations or it is possible that no constitutive models/equations are required. This generic model structure (adapted from Cameron & Gani, 2011) is summarized in Table 2.1.

Table 2.1 *Generic model structure for models in chemical and biochemical engineering (adapted form Cameron & Gani, 2011)*

equation type	***generic form of equations***
balance equations:	$LHS = f(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{p}, \boldsymbol{u}, t, \boldsymbol{\theta}, \boldsymbol{x}), \quad LHS = \begin{cases} 0 \ or, \\ d\boldsymbol{y}/dt \ or, \\ \partial \boldsymbol{y}/\partial t + \sum_i \partial \boldsymbol{y}/\partial u_i \end{cases}$
constitutive equations:	$0 = \boldsymbol{\theta} - g_1(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{p}, \boldsymbol{\theta}, \boldsymbol{x})$
conditional equations:	$0 = \boldsymbol{x} - h(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{p}, \boldsymbol{x})$

In Table 2.1, LHS stands for 'Left Hand Side', $\boldsymbol{y}$ represents the vector of state variables, $\boldsymbol{z}$ represents the vector of known variables and $\boldsymbol{p}$ the vector of model parameters. The variable t represents the independent variable time whereas $\boldsymbol{u}$ is a vector formed by all remaining independent variables considered for the system. Finally, $\boldsymbol{\theta}$ is a vector of constitutive variables which are calculated by the constitutive equations and $\boldsymbol{x}$ is the vector of conditional variables resulting from the conditional equations. Both, $\boldsymbol{\theta}$ and $\boldsymbol{x}$ can either be explicit or implicit variables.

The modelling methodology proposed in this work divides the process of model development into four different phases (see Figure 2.1) associated with detailed work-flows. The different phases are explained below together with the motivation for this particular decomposition.

Phase I: The goal here is to establish the modelling objective and to collect the system information that is required for model development, application and re-use. Different alternative model-scenarios are derived based on the modelling objective and the collected system information.

Phase II: The model construction (single-scale or multi-scale) phase is always needed when the required model is not available in any model library. The goal of this phase is to obtain an implemented and analysed model that is ready-to-use. Therefore, a numerical (model) analysis is also performed in Phase II which includes a degree of freedom analysis, selection of variables to specify, derivation of a solution strategy and link to a required solver.

Phase III: The model identification/discrimination phase is not necessarily required for model development but is only needed if the model contains unknown parameters and adequate experimental data is available/ accessible. In this case the model parameters are regressed (identified).

Phase IV: All model candidates derived in Phase I and developed in Phases II and III are compared, evaluated and validated in this phase. The final optimal model with respect to the modelling goal is selected. In case the performance of none of the candidate models is satisfactory the modeller goes back to the previous phases (Phases I, II, or III) and iteratively improves the model. Note, however, that not all phases and steps of the methodology are

obligatory and their actual use depends on the specific modelling problem and the state of an already existing model. The modeller needs to decide on a case by case basis if a step is relevant for the specific problem and modelling purpose.

In Phase V, the methodology provides two work-flows for model application for: A) Simulation and B) Optimization (see Figure 2.1). If the model results are found to be inadequate for the desired application, the modeller needs to return to the iterative model development process (Phases I-IV), adjust the modelling objective (Phase I) and derive a new more appropriate model.

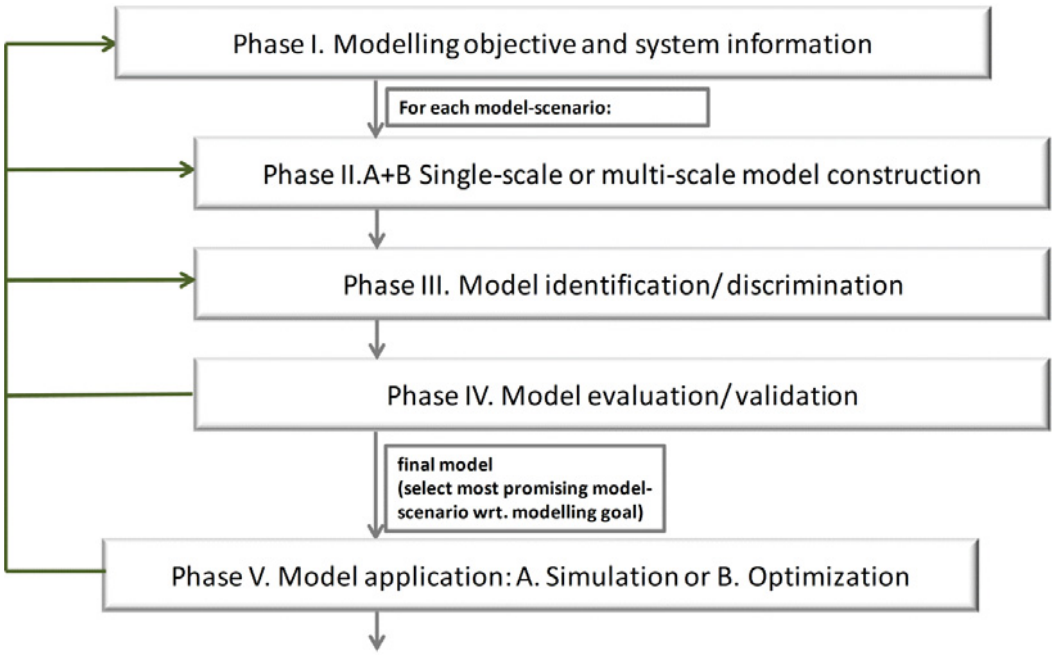


Figure 2.1 *Generic work-flow-based modelling methodology: Overview of generic work-flows for the different modelling tasks and their interconnection.*

There are three different sets of inputs and corresponding outputs for the modelling methodology. The first set is valid for the model development process (Phases I-IV) whereas the second and third sets belong to the work-flows for model application for simulation and optimization, respectively. All three sets of inputs and main outputs are summarized in Table 2.2.

Table 2.2 *Overview of in- and outputs for modelling methodology*

	inputs	***outputs***
Model development (Phases I-IV)	Modelling objective, system information, variable classification (partly), required variable values	Model object ready to apply for simulation or optimization
A. Simulation (Phase V)	Variable values, solver options	Simulation results (plots and tables of variable values)
B. Optimization (Phase V)	Variable values, initial design variables, solver and optimizer options, objective function	Optimized design variables, value of objective function, model/system performance with optimized design variables

Chapter 2 is structured according to the phases of the modelling process and model application (see Figure 2.1). For each phase the corresponding work-flows and data-flows are presented in detail.

2.1 Model development

In this section the different work-flows of the model development process are presented.

2.1.1 Phase I. Modelling objective and system information

The work-flow for the 'Modelling objective and systems information'-phase has the goal to establish the modelling objective as well as to systematically collect and document all information required for the model development and application as well as model re-use by the developer or a different user. Figure 2.2 shows the work-flow and data-flow for the 'Modelling objective and system information'-phase. The different work-flow steps and sub-steps are given in the boxes on the left hand side whereas the data-flow for each step is given in the arrow boxes on the right hand side. For this work-flow and all following work-flows information obtained in one work-flow step is transferred to all following work-flow steps. The arrows on the left hand side indicate the iterative nature of the 'System information and documentation'-work-flow.

The inputs for the work-flow are the modelling objective and the system information. The system information is structured as shown in Table 2.4. The outputs of the work-flow are a structured documentation of all information available on the system to be modelled and possible model-scenarios to be developed applying the work-flows of Phases II-IV.

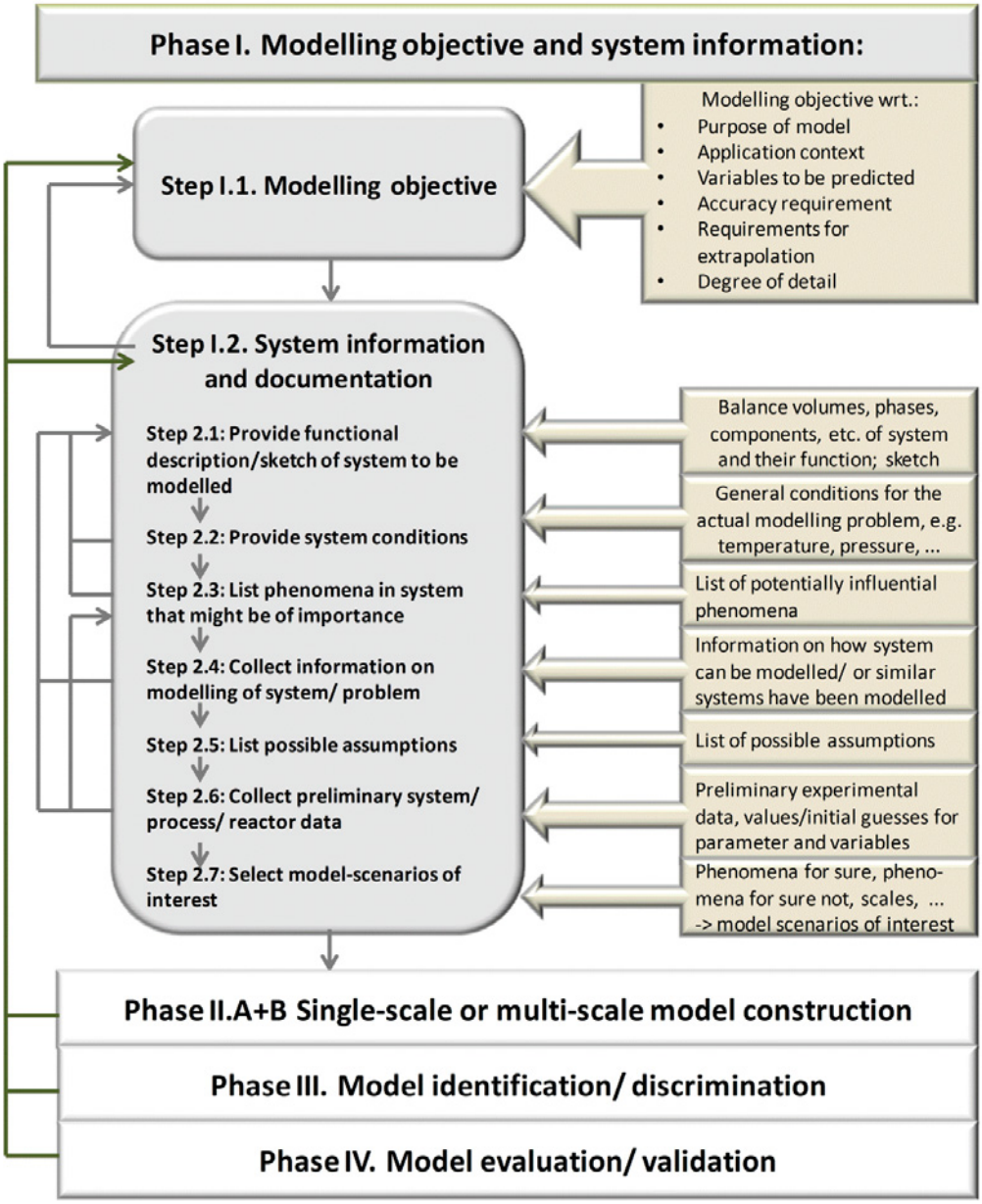


Figure 2.2 *Work-flow (boxes on left hand side) and data-flow (arrow boxes on right hand side) for Phase I: Modelling objective and system information (arrows on left indicate iterative nature of Phase I).*

For each work-flow step the corresponding algorithm describes the objective, the data (required input data), the action (what needs to be done in the step?), rules that may be needed, assumptions made and conditions (e.g. under which circumstances is it necessary to go back to a previous work-flow step?). Not always all mentioned categories are applicable for each work-flow step.

Step I.1. Modelling Objective

Objective: Establish modelling objective with respect to model performance.

Note: These objectives will be used for model performance evaluation and for checking if model reformulations are justified (in Phase IV, see Section 2.1.4).

Action-1: Provide/refine modelling objective.

Note: Table 2.3 summarizes important questions that help to define the modelling objective.

Table 2.3 *Important questions to clarify modelling objective.*

Information included in modelling objective:
Purpose of the model, why is it necessary?, what is it going to be used for?, model application context?
Variables to predict?
Required prediction accuracy?
Requirements for extrapolation?
Degree of detail required (e.g. lumped, steady state, phenomena desired)?
Available/accessible experimental data?

Action-2: Prepare documentation of modelling objective.

Step I.2. System Information and documentation

Objective: Systematic collection and generation of required information on the system under investigation (product, unit operation, process, etc.) to develop, identify, validate and/or extend the model.

The collected information is structured according to Table 2.4 which is filled out in the sub-steps I.2.1-2.7. Possible sources for information are literature, model libraries, databases, experience, expert knowledge and experiments.

Table 2.4 *Structure for system information.*

Categories for system information:
2.1 Functional description/ sketch of the system to be modelled
2.2 System conditions
2.3 Phenomena in system that might be of importance
2.4 Modelling of system/ problem
2.5 Possible assumptions
2.6 Preliminary system data
2.7 Model-scenarios of interest

Step I.2.1 Provide functional description/ sketch of the system to be modelled (process, unit operation, product)

Objective: Functional description and sketch of system in order to get a more detailed overview of the system to be modelled and for later re-use of the model.

Note: The functional description describes the function of the overall system and its sub-units (e.g. components in a formulation/product, unit operations and subunits like membranes, catalyst, different phases). The sketch contains all sub-units of the system with their boundaries and contact areas and additional information including possible phenomena and conditions like temperature and pressure.

Action-1: Write/update functional description of system and prepare/update sketch of system.

Action-2: Prepare documentation.

Step I.2.2 Provide system conditions

Objective: Collection and documentation of relevant system conditions for model application.

Data: System conditions, e.g. (initial) temperatures, (initial) concentrations, (initial) pressures, in-and output streams, chemical compounds.

Action-1: Collect relevant system conditions for modelling problem.

Action-2: Prepare documentation.

Step I.2.3 List phenomena in the system that might be of importance

Objective: List and documentation of all potentially influential phenomena in order to get a more detailed idea on how the system can be modelled.

Action-1: Derive/update list of phenomena that might be of importance for modelling of system and modelling objective.

Action-2: Prepare documentation.

Note: Table 2.5 gives an overview over phenomena that play a role for chemical engineering applications. This list, however, is not meant to be exhaustive.

The considered phenomena determine the degree of detail of the developed model-scenarios and consequently have an impact on the scales to be considered.

Table 2.5 *Overview over important phenomena in chemical engineering.*

phenomenon	***types***
transport	*mass:* convection, diffusion (Fick, Knudsen, surface, Stephan-Maxwell, etc.), dispersion *heat:* convection, conduction, dispersion *impulse*
phase change	solid ↔ liquid, liquid ↔ gas, solid ↔ gas
heating, cooling	
reaction	
adsorption/desorption	
absorption/desorption	
friction	
agglomeration	
breakage	
mixing	
gravity	

Step I.2.4 Collect information on modelling of system/ problem

Objective: Collection and documentation of all available information on how the system can be modelled.

Action-1: Collect all information (if available) on how the system can be modelled/has been modelled before and how similar systems have been modelled. The research for information is done at different levels:

1. Conservation equations and included phenomena;
2. Constitutive equations/ models for different phenomena;
3. Constitutive equations/ models for required properties.

Action-2: Prepare documentation.

Step I.2.5 List possible assumptions

Objective: Documentation of possible assumptions in order to get a more detailed idea on how the system can be modelled.

Note: This list, together with the list of potentially important phenomena support the modeller or a new user of the model later in the iterative model-scenario development to increase or decrease the degree of detail. It moreover provides guidance for a possible adjustment of the model to another application field.

Action-1: Derive/update list of possible assumptions.

Action-2: Prepare documentation.

Step I.2.6 Collect preliminary system/process/reactor data

Objective: Collection and documentation of preliminary data on the system. Establishment of how the desired system behaviour can be monitored quantitatively and what kind of experimental data could be generated for later validation.

Note: The access to experimental data to validate the different parts of a model is one important factor in the decision for the degree of complexity of the final model (see Phase I, Step I.2.7 and Phase IV).

Data: (Preliminary) system data, e.g. experimental data, (initial) parameter and variable values.

Action-1: Collect and update (e.g. after new experiments) available data on system.

Action-2: Prepare documentation.

Step I.2.7 Select model-scenarios of interest

Objective: Derive model-scenarios to be developed.

Action-1: Identify appropriate degree of detail (based on modelling objective and collected system information) and conceptualize corresponding model-scenario(s) to be developed in Phases II-IV using the factors given in Table 2.6.

Table 2.6 *Overview of degree-of-detail determining factors*

degree-of-detail determining factors:
▪ Modelling objective: Required input and output variables/ where is the model going to be applied?/ minimum required degree of detail
▪ Phenomena that might be of importance
▪ Assumptions
▪ Experimental data: a) on which scale/ degree of detail? (needs to be included), b) which scales/phenomena can be identified by available or accessible experimental data?
▪ Geometry and different sub-systems (balance volumes, phases, etc.)
▪ Unknown or too uncertain parameters, or: parameter that change with conditions
▪ Literature/ model libraries/ documentation/ expert knowledge/ experience

Note: Degree of detail has impact on number of scales in model-scenario(s).

Questions to support the derivation of model-scenarios to be developed are:

- Which phenomena occur for sure?
- Which phenomena are important for the modelling goal?

- Which phenomena can be excluded for sure?
- Which is the most influential phenomenon? (potential for model simplification)

Rule-1a: Minimization of modelling errors and effort: Start with a model-scenario which is as simple as reasonable or with model-scenarios already developed (e.g. in literature, model library).

Note: The starting scenario does not necessarily need to consider all phenomena that might be of importance and all scales that have been identified in the preliminary analysis. It is however very important to have these options in mind. Based on the starting scenarios the modeller identifies the optimal final model for the specific modelling goal by extensions and/or simplifications during the iterative modelling procedure (Phases I-IV).

Rule-1b: If modeller does not have any clear idea on how model-scenario looks like the systematic strategies to develop a (multi-scale) model given in Section 1.2 (summarized by Ingram et al. (2004)) are used.

Note: Frequently applied strategies are the top-down or bottom-up approaches. The top-down strategy for example proposes to start the model-scenario development with the highest scale (with respect to modelling goal and desired output variables) and systematically add smaller scales until the desired degree of detail for the modelling purpose is reached. It might turn out at the end that only a single scale is needed.

Action-2: Prepare documentation.

Action-3: Decide to continue with single- or multi-scale model construction work-flow (Phase II, Section 2.1.2).

Rule-3: If modelling scenario of interest is multi-scale or a systematic multi-scale model development strategy is to be applied use multi-scale model construction work-flow (Phase II.B) else use single-scale model construction work-flow (Phase II.A).

Note: It is possible to switch between the work-flows if needed.

2.1.2 Phase II. Model construction

The objective of the model construction work-flows is to systematically and efficiently construct reliable and analysed models for the model-scenarios of interest identified in Phase I based on the modelling objective and system information. Two alternative model construction work-flows have been developed. Phase II.A is dedicated to single-scale model construction whereas Phase II.B provides a work-flow for multi-scale model construction.

Note: For the construction of the models for the different scales within a multi-scale scenario the single-scale model construction work-flow is 'called' for each scale.

II.A Single-scale model construction

The proposed work-flow and data-flow for single-scale model construction is shown in Figure 2.3. The inputs for the single-scale model construction work-flow are the types of balance equations needed and their form, the considered phenomena and assumptions for the phenomena models as well as the classification of the model variables and the required variable values. The outputs of the work-flow are the model equations which are in discretized form if PDEs (Partial Differential Equations) are present, a translated and analysed model together with a solution strategy which is connected to the corresponding solvers – in short, a ready-to-use model object.

Step II.1. Model-Scenario Documentation and Concept

Objective: Establishment and documentation of description of model-scenario based on which the model equations can be derived and for later re-use of the model.

Note: Model description is derived based on modelling objective and system information (Phase I). The upper half of Figure 2.4 summarizes the specifications made in this step.

Action-1: List considered phenomena for each balance volume (balance volumes and phenomena known from Phase I, Steps I.2.1, I.2.7) and assumptions.

Action-2: Identify required conservation equation types (mass, energy, momentum).

Action-3: Decide on form of conservation equations (steady state or dynamic, lumped or distributed, if distributed: in which directions?).

Action-4: Decide on connections between balance volumes.

Action-5: Prepare documentation.

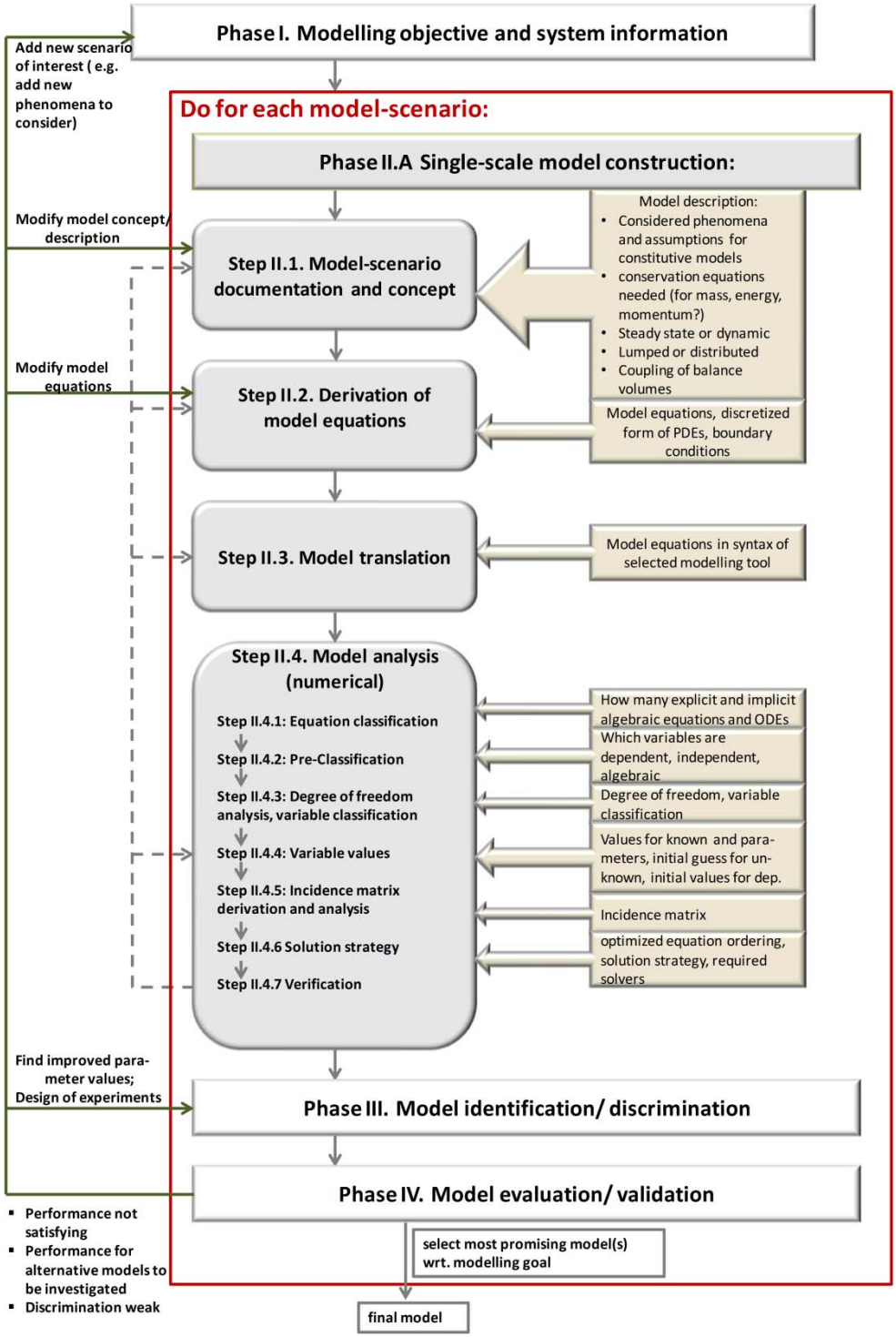


Figure 2.3 *Work-flow (boxes on left hand side) and data-flow (arrow boxes on right hand side) for Phase II.A: Single-scale model construction (arrows on left indicate iterative nature of Phase II.A).*

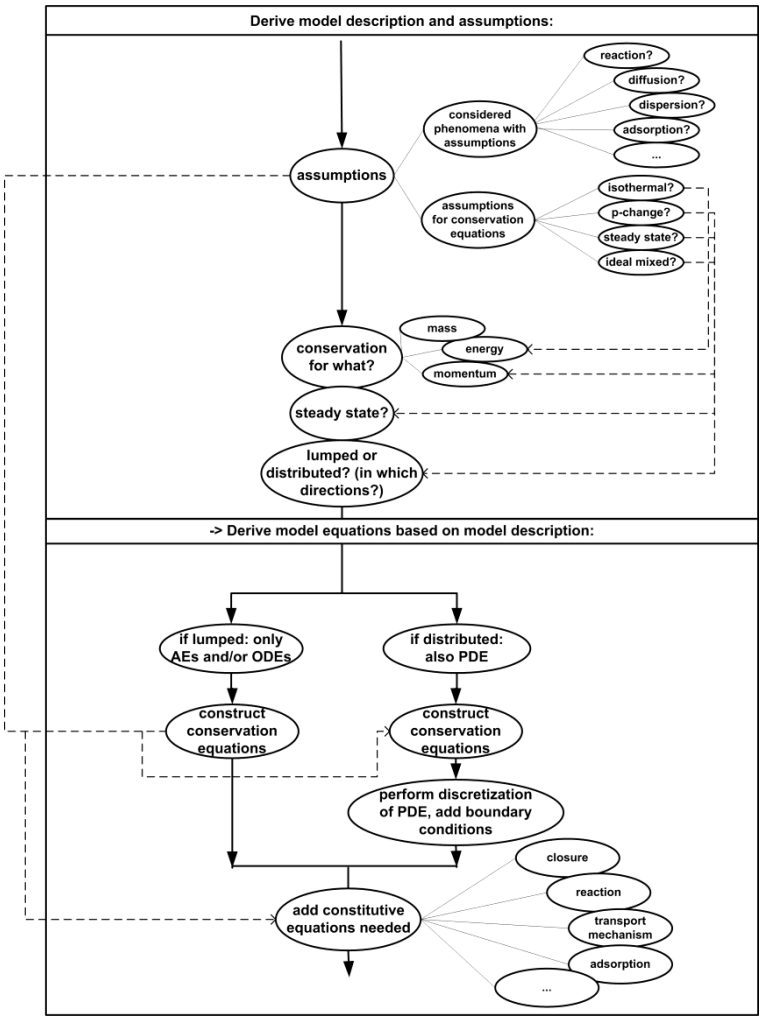


Figure 2.4 *Systematic overview over model descriptions derived in Step II.1 'Scenario documentation and concept' (top) and Step II.2.1 'Derivation of model equations' (bottom).*

Step II.2. Derivation of model equations

Objective: Derivation of model equations for model-scenario.

Note: Model equations are derived based on model description provided in previous step (Step II.1) and information collected in Phase I (Step 2.4) (e.g. model equations for constitutive models). The lower half of Figure 2.4 summarizes the derivation of the model equations.

Action-1: Based on description in Step II.1 (Actions-1, 2, 3): Derive required conservation equations for each balance volume in correct form (steady state or dynamic, lumped or distributed in the different directions) and include terms for considered phenomena.

Note: If the model is lumped the conservation equations are algebraic equations (AEs) or ordinary differential equations (ODEs). If the balance volume however is distributed in one or more coordinate directions the resulting balance equations are partial differential equations (PDEs) or AEs. In the special case of a distribution in only one direction and a steady state model the balance equations can also be ODEs.

Action-2: Discretize models containing PDEs (applying techniques like methods of lines, orthogonal collocation, finite elements, finite volumes) and add corresponding boundary conditions.

Action-3: Add required constitutive models (for considered phenomena, properties, etc.) based on Step II.1 (Action-1) and Step I.2.4.

Action-4: Provide connecting equations for different balance volumes based on description in Step II.1 (Action-4).

Action-5: Prepare documentation.

Note: Additional details on the systematic derivation of model equations are given by Jensen & Gani (1996).

Step II.3 Model translation

Objective: Create translated model object.

Note: In this context 'model translation' means the transformation of the model equations available in text-format to a model object that is readable by the computer and can be connected to the required solvers. Different methods can be applied for model translation, for example, 'Reverse Polish Notation' (RPN).

Action-1: Select modelling tool that allows connection to the required solvers (dynamic, algebraic, optimizer).

Action-2: Enter model equations in corresponding syntax.

Action-3: Translate/compile model.

Action-4: Prepare documentation.

Step II.4. Model Analysis (Numerical)

Objective: Ensure consistency of model with respect to variable types, degree of freedom, singularities, structure. Derivation of solution strategy and verification in general. Identification and connection of required solver(s).

The following sub-steps are identified:

Step II.4.1 Equation classification

Objective: Obtain number and types of equations (needed for following sub-steps).

Action-1: Classify equations according to their type as either algebraic or differential.

Action-2: Prepare documentation.

Step II.4.2 Pre-classification of variables

Objective: Pre-classification of variables according their types as either algebraic, dependent (differential) or independent (needed for following sub-steps).

Action-1: Classify all model variables as either dependent, independent or algebraic.

Action-2: Prepare documentation.

Step 4.3 Degree of freedom analysis

Objective: Satisfy degrees of freedom in order for model to become solvable.

Action-1: Calculate degrees of freedom for algebraic equation part DOF_{AE}.

Rule-1: DOF_{AE}= number of algebraic variables in AE-part – number of AEs.

Action-2: Satisfy DOF_{AE} by specifying required amount of algebraic variables as either known or parameter.

Note: There are three types of known variables: 1. fixed by problem, 2. fixed by system, 3. fixed by model.

Action-3: Calculate degree of freedom for ODE-part of model DOF_{ODE}.

Rule-3: DOF_{ODE} = Number of algebraic variables in ODE part that do not occur in AE part of model.

Action-4: Satisfy DOF_{ODE} by specifying corresponding algebraic variables as known or parameter.

Action-5: Specify all remaining algebraic variables as unknown variables.

Action-6: Prepare documentation.

Step II.4.4 Variable values

Objective: Provide variable values needed for model solution.

Data: Variable values:

- for parameter and known variables: value or initial guess (if they are to be identified or optimized in a later step);
- for unknown variables: initial guess;
- for dependent variables: initial value.

Action-1: Provide variable values.

Action-2: Prepare documentation.

Step II.4.5 Incidence matrix derivation and analysis

Objective: Create and analyse incidence matrix to identify optimal equation ordering and solution strategy.

Action-1: Create incidence matrix by:

1. Write model equations in rows of matrix (2 parts: AE- and ODE-part);
2. Write variables (unknown and dependent) to columns;
3. Mark which variables appear in which equations.

Action-2: Find optimal equation order by transforming AE- and ODE-part of incidence matrix as close as possible to lower tridiagonal form and analyze incidence matrix following Rules-2a-d.

> *Rule-2a:* If ODE/AE parts are of lower tridiagonal form all ODEs/AEs can be solved 1-by-1.
>
> *Rule-2b:* If ODE-part/AE-part of incidence matrix has off-diagonal elements corresponding ODEs/AEs need to be solved together.
>
> *Rule-2c:* If ODE-part contains unknown algebraic variables (from AE-part) AEs need to be solved (until convergence) before solving ODEs.
>
> *Rule-2d:* If AE-part contains dependent variables (from ODE-part) AE- and ODE-parts need to be solved coupled and AEs have to be solved for each time-step of ODEs.

Action-3: Order model equations accordingly.

Action-4: Prepare documentation.

Note: Since the incidence matrix shows which model equations need to be solved together and which are de-coupled, in that way it supports the modeller in decomposing a complex model into sub-models.

Step II.4.6 Solution strategy

Objective: Derive solution strategy, connect and set-up required solvers in order to be able to verify and solve model.

Action-1: Identify required solvers (based on equation and variable types: ODE, AE, DAE, implicit AEs?).

Action-2: Connect solvers to model.

> *Condition-2:* Required solvers cannot be connected to model.
>
> If *Condition-2* do: Change software environment.

Action-3: Perform eigenvalue analysis in order to evaluate stiffness of dynamic models.

> *Rule-3a:* If $\frac{|\lambda_{max}|}{|\lambda_{min}|} > 1000$ (the ratio of the maximum absolute real part and the minimum absolute real part of the eigenvalues of the system's Jacobian matrix) the system is regarded as 'stiff'.

Note: A stiff dynamic system consists of modes which reach steady state very fast and slow modes.

Rule-3b: If system is stiff this indicates that the number of dominant eigenvalues is less than the number of equations. This indicates a potential for model reduction/ simplification.

Condition-3b: Model is to be reduced.

If *Condition-3b* do: Go back to Phase I and create the new simplified model-scenario.

Note: If for example two modes are very fast and two others are very slow and the modeller is interested in the dynamic behaviour towards the overall steady state(s) of the system the fast modes can be assumed to be at their steady state solution from the beginning of the simulation. If on the other hand the modeller is interested in results before even the fast modes reach steady state the change of the slow modes can be neglected.

Rule-3c: A stiff system contains different time-scales. Solve all model equations together at the time-scale dictated by the fastest modes.

In case solution of overall model at lowest time-scale is intolerably slow split up model in two or more sub-models with different time-scale and switch to multi-scale model construction work-flow. Split-up involves certain loss of accuracy of model solution.

Rule-3d: If Jacobian matrix is not of full rank (the number of eigenvalues is less than the number of equations) this indicates that the system contains linear dependent equations. Go back to Phase I and reduce model.

Action-4: Prepare documentation.

Step II.4.7 Verification

Objective: Verification that constructed model including its model equations, variable values and implementation is reflecting system.

Action-1: Singularity check.

Action-2: Units check.

Action-3: Check for typing mistakes in equations and variable values.

Action-4: Equation-by-equation debugging.

Action-5: Compare to similar already verified and validated models (e.g. simpler model-scenarios) -> change to simulation work-flow (Section 2.2).

Action-6: Check if model produces desired output based on desired input.

Action-7: Prepare documentation.

With this last step the single-scale model construction is completed.

Note: Steps II.1-II.4 are repeated if multiple models are to be developed.

If required the modeller continues with Phase III in order to identify unknown model parameter values and/or to obtain measures to discriminate between different candidate models. Otherwise, a model evaluation and validation is performed directly (Phase IV).

II.B Multi-Scale model construction

The proposed work-flow and data-flow for multi-scale model construction is shown in Figure 2.5. The inputs for the multi-scale model construction work-flow are initial information about the considered phenomena, assumptions, required scales and/or sub-models, integration frameworks and data-flow. Furthermore, the variable classification and the required variable values need to be specified. The outputs of the work-flow is a linked, translated and analysed multi-scale model together with a solution strategy which is connected to the corresponding solvers – in short, a ready-to-use model object.

Step II.1. Model-scenario documentation and concept

Objective: Establishment and documentation of description of model-scenario based on which the scenario can be constructed and for later re-use.

Action-1: List/update (initial) information about assumptions, considered phenomena, scales and/or required (sub-)models, integration framework (see Chapter 1), data-flow.

Note: The (initial) scenario concept is derived based on the modelling objective and system information (Phase I).

Action-2: Create/update (preliminary) version of data-flow scheme.

Note: A data-flow represents the different scales in the multi-scale scenario as boxes and indicates the data-flow between the scales by arrays. The communicated variables between scales are written next to the corresponding array. For each scale the output variables are given.

Action-3: Prepare documentation.

Action-4: Repeat following Steps II.2 and II.3 in a loop over all scales in the multi-scale scenario.

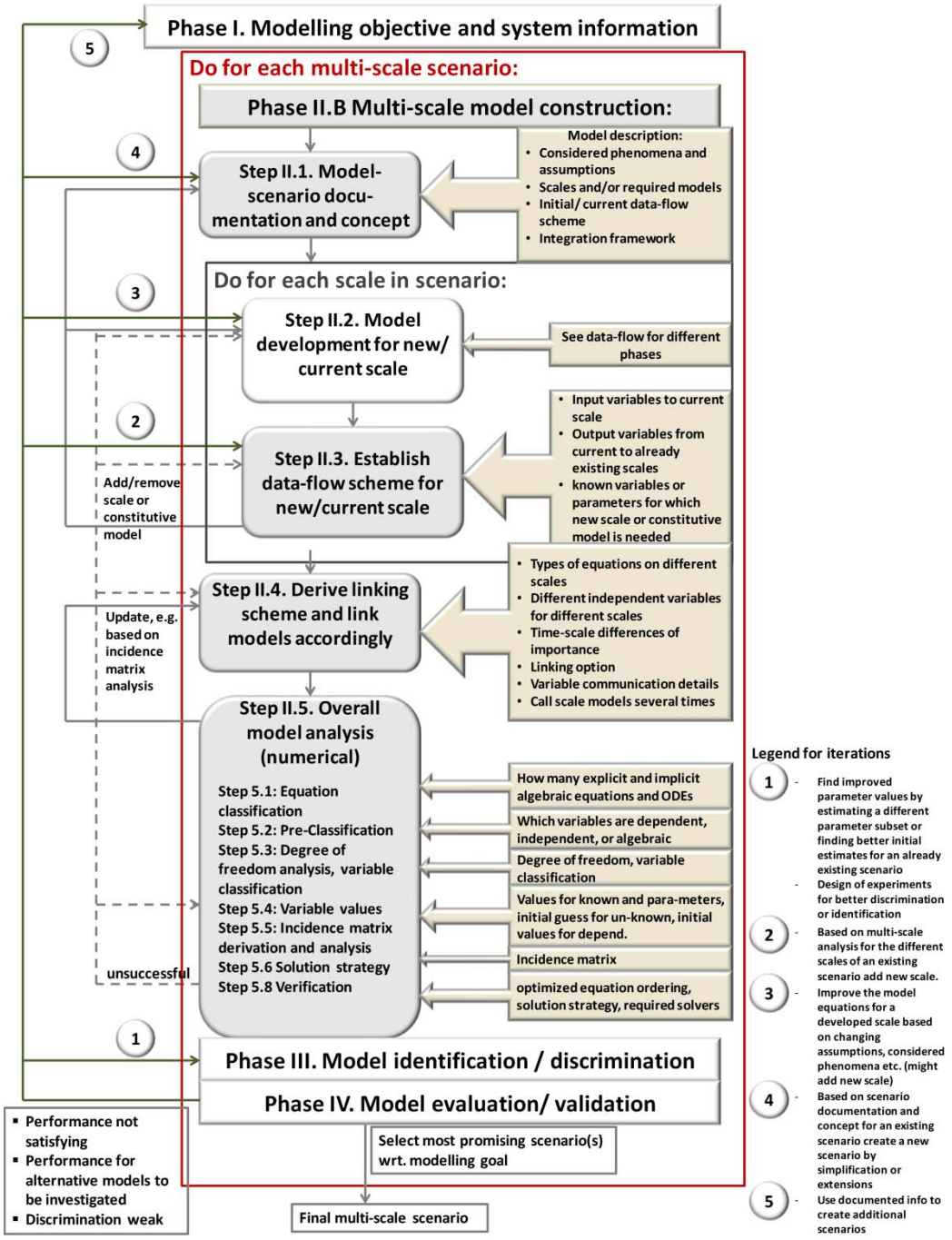


Figure 2.5 *Work-flow (boxes on left hand side) and data-flow (arrow boxes on right hand side) for Phase II.B: Multi-scale model construction (arrows on left indicate how Phase II.B is iterative).*

Step II.2. Model development for new/current scale

Objective: Develop model for current scale.

Data: See data required for Phases I-IV.

Action-1: Go through all phases of model development for current scale (Phases I-IV).

Note: In most cases single-scale model construction work-flow is applied for model construction (Phase II) of the different scale-models. In some cases, however, the model for a scale is also a multi-scale model.

Note: During numerical model analysis of scale-model (in Phase II.A, Step 4): Not necessary to provide variable values for known variables and parameters that are communicated from another scale or sub-model in multi-scale scenario.

> *Condition-1:* Current scale is to be split into one or more different time-scale models (decided based on eigenvalue analysis (Phase II.A, Step 4) during development of model for current scale)
>
> If *Condition-1* do: Go back to Step II.1 (Actions-1-4) and update multi-scale scenario concept.

Action-2: Prepare documentation.

Step II.3. Establish data-flow scheme for new/current scale

Objective: Establish data-flow scheme between current scale and connected scale(s).

The following sub-steps are identified:

Step II.3.1 Update data-flow between existing scales

Objective: Establish data-flow between current scale and connected scale(s).

Action-1: Identify variables in scale-model that are communicated to already developed models of other scales in scenario.

Action-2: Identify variables in scale-model that are communicated from already developed models of other scales in scenario.

Action-3: Prepare documentation.

Step II.3.2 Multi-scale analysis

Objective: Identification of additional scales that can/need to be added to multi-scale scenario.

Action-1: Top-down approach: Check if additional model/ constitutive equation required for one or more of the known variables or parameters in current scale model.

> *Rule-1a:* Start with simplest model-scenario reasonable and evaluate its performance (Phase IV) before going to more complex scenarios.

Rule-1b: Adding of additional model for a 'parameter' or known variable (top-down) useful if:

- No value available for known variable;
- Lack of experimental data on current scale -> Parameter cannot be estimated (correlation or model is used instead);
- Variable value changes with changing conditions (like e.g. temperature) -> Adding a model for this variable makes model predictions more exact and/or better extrapolable to systems with different conditions.

Condition-1: Additional model needed.

If *Condition-1* do: If additional model is on same scale, go back to Step II.2 (Actions-1-2) to adjust model assumptions and equations, else go back to Step II.1 (Actions-1-4) and update scenario concept.

Action-2: Bottom-up approach: Check if model/scenario is to be applied in a higher scale model.

Condition-2: Additional model needed.

If *Condition-2* do: If additional model is on same scale, go back to Step II.2 (Actions-1-2) to adjust model assumptions and equations, else go back to Step II.1 (Actions-1-4) and update scenario concept.

Action-3: Prepare documentation.

Step II.4. Derive linking scheme and link models accordingly

Objective: Establish linking scheme between scales in multi-scale scenario and link models accordingly to obtain model object.

Note: Linking-scheme is constructed based on data-flow scheme established in previous step (Step II.3) and integration framework (see Chapter 1).

Action-1: Identify types of equations on the different scales.

Action-2: Identify independent variables for each scale that contains ODEs. If more than one scale within linking scheme consists of dynamic models: Identify time-scale difference between scales.

Action-3: Decide for linking options for the different scales (Option 1: Combine all model equations as one set of model equations, Option 2: Call different models as sub-models according to linking scheme).

Rule-3: Choose Option 2 for linking two scales only if a) time-scales between linked scales differ significantly and combined numerical solution not feasible (e.g. high computational effort), b) different independent variables for linked scales.

Note: For Option 1 model solution is more exact. Option 2 allows solving each model at its time-scale instead of solving all linked models at the time-scale dictated by the fastest modes in the overall system.

Action-4: Specify number of calls/solution steps (under different conditions) for each scale-model in linking scheme.

Action-5: Provide options for variable communication:

- Multiply variable by a constant factor (e.g. multiply by number of catalyst particles in ideal mixed bulk);
- Form average value of a vector/matrix variable before communication to other scale;
- Form average value of an output variable calculated in multiple calls of scale model before communicating it to other scale;
- Sum-up elements of a vector/matrix variable before communication
- Sum-up values of an output variable calculated in multiple calls before communicating it to other scale;
- Communicate boundary values/1 element of a vector variable.

Note: Possible reasons for solving a scale model several times under different conditions are for example:

- Calling scale model is discretized;
- Calling scale model contains a population balance that requires solving the lower scale model several times for example for different particle diameters.

Possible reasons for communicating average value:

- Scale-model is called several times (not always necessary for this case).
- At least one of the connected models is discretised.

For the case of parallel coupling of two discretised models several average values need to be communicated from the finer distributed model, one for each discretisation element of the coarser model.

Action-6: Adjust vector/matrix dimensions of communicated variables in connected scale-models.

Action-7: Derive solution sequence of scale-models based on data-flow (solution sequence is optimized/ confirmed based on the incidence matrix derived in the following numerical model analysis step (Step II.5)).

Action-8: Derive information on which scales are to be solved integrated based on data-flow (also this information can be systematically obtained based on incidence matrix in Step II.5).

Action-9: Link models according to made specifications and add linking equations according to specifications made for variable communications (e.g. averaging equations).

Action-10: Prepare documentation.

Step II.5. Overall model analysis (numerical)

Objective: Ensure consistency of overall multi-scale model with respect to variable types, degrees of freedom, singularities, structure. Derivation of solution strategy, connection of required solvers and verification in general.

Action: Perform numerical model analysis of overall multi-scale model following the algorithm given in Step II.4 of the single-scale model construction work-flow (Phase II.A).

Note:. For eigenvalue analysis the Jacobian matrix of the overall model, or at least all coupled differential equations, has to be considered because the eigenvalues of a coupled system differ from the eigenvalues obtained from the sub-systems.

With this last analysis step the multi-scale model construction is completed.

Note: Steps II.1-II.5 are repeated if multiple multi-scale models are to be developed.

If needed the modeller applies the 'Model identification/discrimination'-work-flow (Phase III) in order to estimate unknown model parameters and/or discriminate between different candidate (multi-scale) model scenarios. Afterwards, an evaluation and validation of the constructed (multi-scale) model scenarios is performed (Phase IV).

2.1.3 Phase III. Model Identification/ discrimination

The work-flow for the 'Model identification/ discrimination'-phase has the goal to identify the model parameters and to discriminate between different candidate models based on experimental data. Figure 2.6 shows the work-flow and its integration in the overall modelling methodology.

The inputs for the work-flow are the experimental data, the list of parameters to be identified, initial values for these parameters, the type of the objective function or discrimination measure and applied optimization algorithm. The outputs are the estimated values of the parameters including statistics (e.g. sensitivities, confidence intervals) and the value of the objective function or discrimination measure.

Note: There are no separate work-flows for model identification and discrimination because these tasks are strongly related in the sense that the identification of all model candidates and the subsequent comparison based on the identification results (e.g. value of objective function and confidence intervals) is nothing else than model discrimination. Consequently, whenever there are more than one candidate model to be identified the problem is a model discrimination problem.

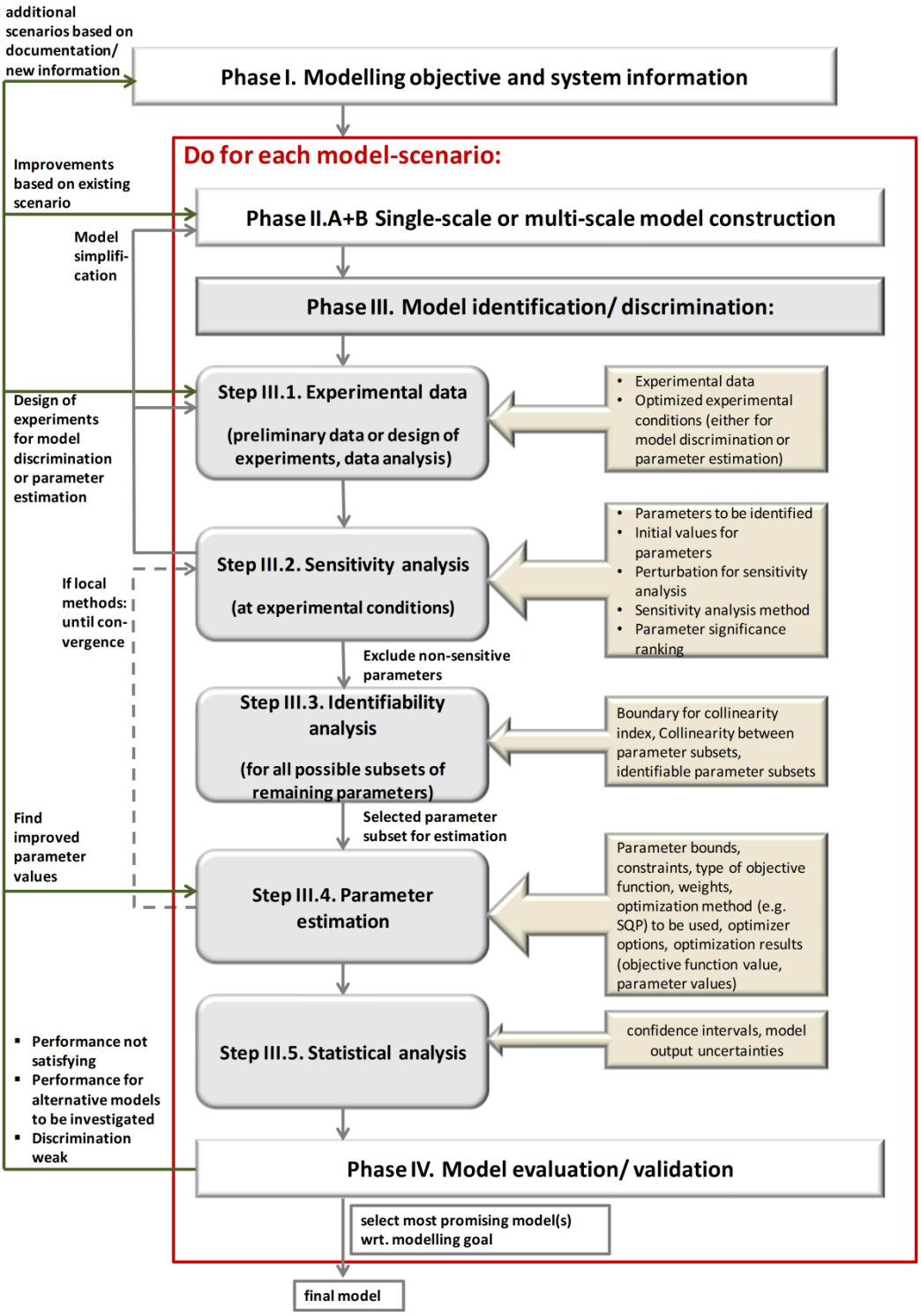


Figure 2.6 *Work-flow (boxes on left hand side) and data-flow (arrow boxes on right hand side) for Phase III: Model identification/discrimination (arrows on left indicate how Phase III is iterative).*

Step III.1. Experimental data

Objective: Analysis of available experimental data (collected and documented in Phase I). Design new experiments to increase identifiability of unknown model parameters (decrease confidence intervals) or improve discrimination between two or more candidate model-scenarios.

Data: Documented experimental data in Phase I or new data.

Condition-1: No experimental data available or dedicated experimental data for the specific model identification/ discrimination problem required.

If *Condition-1*go to *Action-1*.

Action-1: Perform design of experiments and corresponding experiments.

Note: Detailed work-flows/algorithms for design of experiments not included in methodology.

Note: In many cases, the experimental data required cannot be generated. This can be due to non-measurable variables, not realizable measurement conditions, inexact measurement techniques or time and cost constraints. In these cases the model needs to be revised (simplified) or the modeller has to accept the uncertain parameter estimates and blurred discrimination.

Condition-2: Available experimental data is to be used for model identification/ discrimination.

If *Condition-2* go to *Action-2*.

Action-2 Perform data analysis.

Note: Detailed work-flows/algorithms for data analysis not included in methodology.

Action-3: Prepare documentation.

Step III.2. Sensitivity analysis

Objective: Identification of parameters that have no impact on measured variables under experimental conditions -> decrease size of parameter estimation problem and obtain meaningful parameter estimates.

Note: Sensitivity of experimental data with respect to a parameter at experimental conditions is a necessary condition for identifiability of the parameter. Low sensitivity results in high confidence intervals of the parameter estimates. If measured variables are non-sensitive to all unknown model parameters the representation of the experimental data by the model will not be significantly improved during the model identification phase. By identifying parameters with low sensitivity which can be omitted in the parameter estimation problem this step has the potential to reduce the size of the parameter estimation problem.

Note: Sensitivity analysis is not an obligatory step for model identification/ discrimination. If sensitivity analysis is omitted it is not possible to reduce the size of the parameter estimation

problem and all unknown parameters need to be estimated. The confidence interval calculation performed after parameter estimation provides a measure for identifiability of each parameter but this does not resolve if identifiability issues result from low sensitivity or collinearity between parameters.

Action-1: Decide between local or global sensitivity analysis and for method.

Note: In general, global sensitivity analysis methods give a more complete picture because they do not only perturb one parameter at a time for the same percentage but instead perform a large number of sensitivity runs perturbing all parameters at the same time for differing percentages. Which parameters are perturbed by which percentage for each sensitivity run is determined by a sampling method. The disadvantage of the global methods is that they are computational more demanding than local methods.

Action-2: Connect implementation of chosen method with model.

> *Condition-2:* Software environment does not support sensitivity analysis or connection to external software/implementation of sensitivity analysis
>
> If *Condition-2* do: a) change software environment, b) perform simulations needed for sensitivity analysis 1-by-1 manually (time-consuming), or c) omit sensitivity analysis for parameter estimation.

Action-3: Set unknown model parameters to perturbation variables for sensitivity analysis.

Action-4: Provide initial guesses for parameters.

Action-5: Give boundaries for perturbation of parameters.

Action-6: Set up multiple sensitivity analysis runs for conditions of each experimental data point available.

> *Condition-6:* Set-up of multiple sensitivity runs not supported by model environment/ sensitivity analysis implementation.
>
> If *Condition-6* do: Perform multiple runs manually.

Action-7: Adjust solver selection and solver options for each run to experimental conditions (e.g. output times of dependent variable values).

Action-8: Assign measured variables as response variables for sensitivity analysis.

Action-9: Perform sensitivity calculations (for each response variable and conditions at each experimental data-point).

Action-10: Evaluate obtained parameter significance ranking.

> *Condition-10:* Parameters with low sensitivity (compared to other parameters).
>
> If *Condition-10* do: a) fix non-sensitive parameters to their initial values, or b) design of experiments for parameter estimation (go back to Step III.1), or c) simplify model by removing or lumping non-sensitive parameters (go back to Phase II).

Note: With respect to model simplifications due to non-sensitive parameters it has to be kept in mind that a parameter that has turned out to be insensitive during this analysis is not necessarily insensitive for all model outputs under all possible conditions but just for the measured variables under the experimental conditions.

Rule-10: Do not simplify model due to non-sensitive parameter before further analysis if it is to be applied for the prediction of other variables than the measured variables and other conditions than the experimental conditions

Action-11: Prepare documentation.

Step III.3. Identifiability analysis

Objective: Identification of parameter subsets that are non-collinear and therefore identifiable -> decrease size of parameter estimation problem and obtain meaningful parameter estimates.

Note: Collinearity between parameters results in high confidence intervals. In summary, sensitivity analysis and identifiability analysis ensure the quality of the parameter estimates, identify which parameters cannot be estimated by the available data and thereby also reduce the size of the optimization problem to be solved in the next step.

Note: Identifiability analysis is not an obligatory step for model identification/ discrimination. If identifiability analysis is omitted it is not possible to reduce the size of the parameter estimation problem due to collinearities. The confidence interval calculation performed after parameter estimation provides measure for identifiability of each parameter but this does not resolve if identifiability issues result from low sensitivity or collinearity between parameters and also gives no information on non-collinear subsets.

Note: The identifiability analysis is conducted based on the sensitivity analysis results (sensitivity sub-matrix of unknown parameters that have been deemed sensitive).

Action-1: Connect implementation of collinearity index calculation with model.

Condition-1: Software environment does not support generation of all possible parameter subsets and/or calculation of collinearity index.

If *Condition-1* do: a) change software environment, or b) write external program to do that (needs parameters and sensitivity matrix as inputs), or c) omit identifiability analysis for parameter estimation.

Action-2: Generate all possible subsets of unknown model parameter to estimate.

Action-3: Calculate collinearity index for all subsets (Brun et al., 2002; Sin & Vanrolleghem, 2007).

Action-4: Evaluate, based on collinearity index, which parameter subsets are promising for identification and select final parameter (sub)set(s) for parameter estimation.

Rule-4a: A parameter subset is not identifiable if its collinearity index exceeds threshold (between 10 and 20 (Brun et al, 2002; Sin & Vanrolleghem, 2007)).

Rule-4b: Start by selecting the parameter subset which has the lowest collinearity index from the group of identifiable parameter subsets with the largest number of parameters.

Condition-4: Not all unknown model parameters to be identified appear in final parameter subset.

If *Condition-4* do: For unknown parameters not in final subset: a) fix parameters to their initial values, or b) design of experiments for parameter estimation (go back to Step III.1), or c) simplify model by removing or lumping collinear parameters (go back to Phase II).

Note: With respect to model simplifications it has to be kept in mind that a parameter subset that has turned out to be collinear during this analysis is not necessarily collinear for all model outputs under all possible conditions but just for the measured variables under the experimental conditions.

Rule-4c: Do not simplify model due to collinearities between parameters before further analysis if model is to be applied for the prediction of other variables than the measured variables and other conditions than the experimental conditions

Action-5: Prepare documentation.

Step III.4. Parameter estimation

Objective: Estimate values for unknown and identifiable model parameters by available experimental data.

Action-1: Provide/update initial values for identifiable parameters.

Action-2: Provide boundaries for identifiable parameters.

Action-3: Put constraints on other model variables if applicable.

Action-4: Select objective function (e.g. least square fit, maximum likelihood) including possible weight factors and model for experimental error (if needed for objective function)).

Rule-4: In case a certain model discrimination method, which for example includes a penalty on the number of parameters, is to be used to compare different candidate models in the upcoming 'Model evaluation/validation'-phase (Phase IV) use corresponding objective function during this parameter estimation step.

Action-5: Decide for optimization algorithm (e.g. SQP).

Action-6: Create optimization problem with goal to minimize objective function subject to equality (model equations) and inequality constraints (parameter bounds, constraints on other model variables):

- Connect implementation of optimizer based on selected optimization algorithm to model and its numerical solver;
- Transform model equations, parameter boundaries and specified constraints on other model variables to constraints of optimization problem;
- Implement and connect objective function calculation to optimization problem.

Condition-6: Applied modelling tool does not allow creation of optimization problem.

If *Condition-6* do: Change tool.

Action-7: Set up optimizer options (e.g. convergence tolerance).

Action-8: Set up or implement desired outputs (e.g. variable plots and/or values, at least: estimated parameter values and value of objective function).

Action-9: Perform parameter estimation.

Action-10: Prepare documentation.

Step III.5. Statistical analysis

Objective: Access the model prediction quality based on the confidence intervals of the estimated parameters and induced uncertainties on the model predictions resulting from a statistical model analysis. In that way, the modeller is aware of the extent of the uncertainties in the model predictions that originate from the model development process.

Note: This is not an obligatory step for model identification/ discrimination.

Action-1: Calculate confidence intervals of parameter estimates.

Action-2: Perform uncertainty analysis for model outputs based on obtained confidence intervals.

Action-3: Evaluate results and decide if confidence intervals and uncertainties are acceptable.

Note: Performance criteria: differs with respect to the desired application and modelling goal.

Condition-3: Confidence intervals and/or resulting uncertainties are too high.

If *Condition-3* do: a) Go back to Steps III.2 and III.3 and decide for different parameter subset to estimate, b) Go back to Step III.1 and perform design of experiments for parameter estimation, c) Simplify model based on results obtained in Steps III.2 and III.3.

Action-4: Prepare documentation.

2.1.4 Phase IV. Model evaluation/ validation

For this step no detailed work-flow has been developed but the main points are summarized.

Model Evaluation

The different model-scenarios are compared in order to select the best option with respect to the modelling objective. The criteria for the comparison and evaluation of the different candidate models have been calculated during the 'model identification/ discrimination' phase (Phase III). These are:

- Value of objective function/ discrimination measure;
- Plot of experimental data compared to simulation results;
- Confidence intervals;
- Results from uncertainty analysis;

Not always all these criteria are evaluated and especially the uncertainty results do not need to be created in any case and only for the most promising or the selected final scenario. For comparison between more than one model-scenario a model discrimination criteria which in addition to the objective function value also includes a penalty on the number of model parameters can be employed in case the corresponding objective function has been chosen during the parameter estimation in Phase III. If for example a maximum likelihood objective function has been used the BIC measure can be calculated in this step for the different model candidates. In order to calculate these measures the number of parameters must be set to the number of actual unknown model parameters and not to the number of parameters in a possibly reduced parameter set (based on sensitivity and identifiability analysis in Phase III) used during model identification. This is due to the fact that the model still contains all the parameters. In addition, however, the criteria can be calculated using a parameter number which is reduced by the non-identifiable parameters. In this way the maximum potential of improving the model discrimination-score for a certain model due to simplification based on elimination of non-identifiable parameters can be evaluated. If the potential is big the modeller can consider a simplification of the model but only after confirming the irrelevance of the non-identifiable parameters for all conditions and response variables related to the modelling goal. In addition to the above described criteria for model discrimination other considerations like a tolerable demand of computing-power are also of importance. Furthermore, the modeller needs to evaluate if all phenomena and/or scale sub-models within the overall model scenario can be validated by experimental data. If this is not the case often a simplification of the model is advisable. However, this depends a lot on the modelling goal in terms of the extrapolation requirements and desired system knowledge gain. If no experimental data is available the performance of a model-scenario can be compared to a more complex model.

If the performance of none of the developed scenarios is satisfying with respect to the modelling goal or if the performance of an alternative scenario should be investigated (simplified or of increased degree of detail) the modeller has the following options to iterate on the model performance:

1. Go back to Phase III and obtain improved parameter estimates (for example by finding better initial estimates or estimating other parameter subsets) or perform design of experiments to obtain better quality of parameter values (decrease confidence intervals) and/or model discrimination. The new data needs to be documented (Phase I).
2. Go back to Phase II and improve the models for the scales and considered phenomena of an existing model-scenario (e.g. modify concept of model scenario, model equations, discretization, linking options). This might add new scales.
3. Go back to Phase I and add a new scenario of interest based on documented information (for example by considering additional phenomena) Here, the re-consideration of the different established degree-of-detail-factors can provide new options (see Phase I, Step I.2.7). It might be necessary to generate additional system information.

In order to identify potential for improving a model scenario it is advisable to perform a sensitivity analysis with respect to known variables and parameter values. This will give an idea of how variations in these variables will change the model predictions. If the model outputs turn out to be very sensitive to a variable that has been fixed to a constant estimate but in reality changes with changing conditions there is potential to improve the model predictions by replacing the constant estimate by a model (see protein uptake case study, Section 5.2). This analysis can be especially attractive if the modeller starts from a very simple (multi-scale) scenario to systematically decide how and if the degree of detail should be increased.

Furthermore, for model improvement it is especially helpful to consider and update the factors determining the required degree of detail which have been identified in Phase I. In general it is not always improving the model if the degree of detail is increased (by adding new scales, constitutive equations) because the potential for improvement depends heavily on the performance of the added models. However, the more complex a scenario the better are the extrapolation qualities to other system conditions.

Model validation

It is good practice to validate the final model by independent experimental data which has not been used during the model building stage.

The resulting identified and validated model can then be used for the engineering purpose it was built for like performing simulations to predict system behaviour (follow simulation work-flow) or solving an optimization problem (follow optimization work-flow).

2.2 Model Application

In this section two model application work-flows for simulation and optimization are presented.

2.2.1 A. Simulation

The model application work-flow for simulation has the objective to predict the behaviour of the modelled system under certain conditions. Figure 2.7 shows the proposed work-flow for simulation and its integration with the work-flows of the model development process.

The simulation work-flow requires the input of the simulation objective as well as an update of the variable values according to the conditions to be simulated. For an algebraic system the outputs are the values of the unknown variables. For a differential system profiles of the dependent and unknown variables with respect to the independent variables are calculated. In addition, the possible steady states, their stability and the stiffness of the system can be obtained.

Step A.1. Simulation objective

Objective: Establish simulation objective -> required degree of detail of model including variables to be predicted and their accuracy.

Note: The simulation objective focuses on the model application and what has to be predicted. It can for example be related to monitoring the system functionality and checking how the system performs under given conditions.

Action-1: Formulate simulation objective.

> *Condition-1:* Available model not applicable for specific simulation problem.
>
> **Note:** Model might not have been specifically developed for this simulation problem (e.g. different chemical compounds, kinetic model, temperature, degree of detail, etc.).
>
> If *Condition-1* do: Go back to model development process, adjust modelling objective (Phase I, Step I.1) to application needs and based on that go through Phases I-IV to develop appropriate model.

Action-2: Prepare documentation.

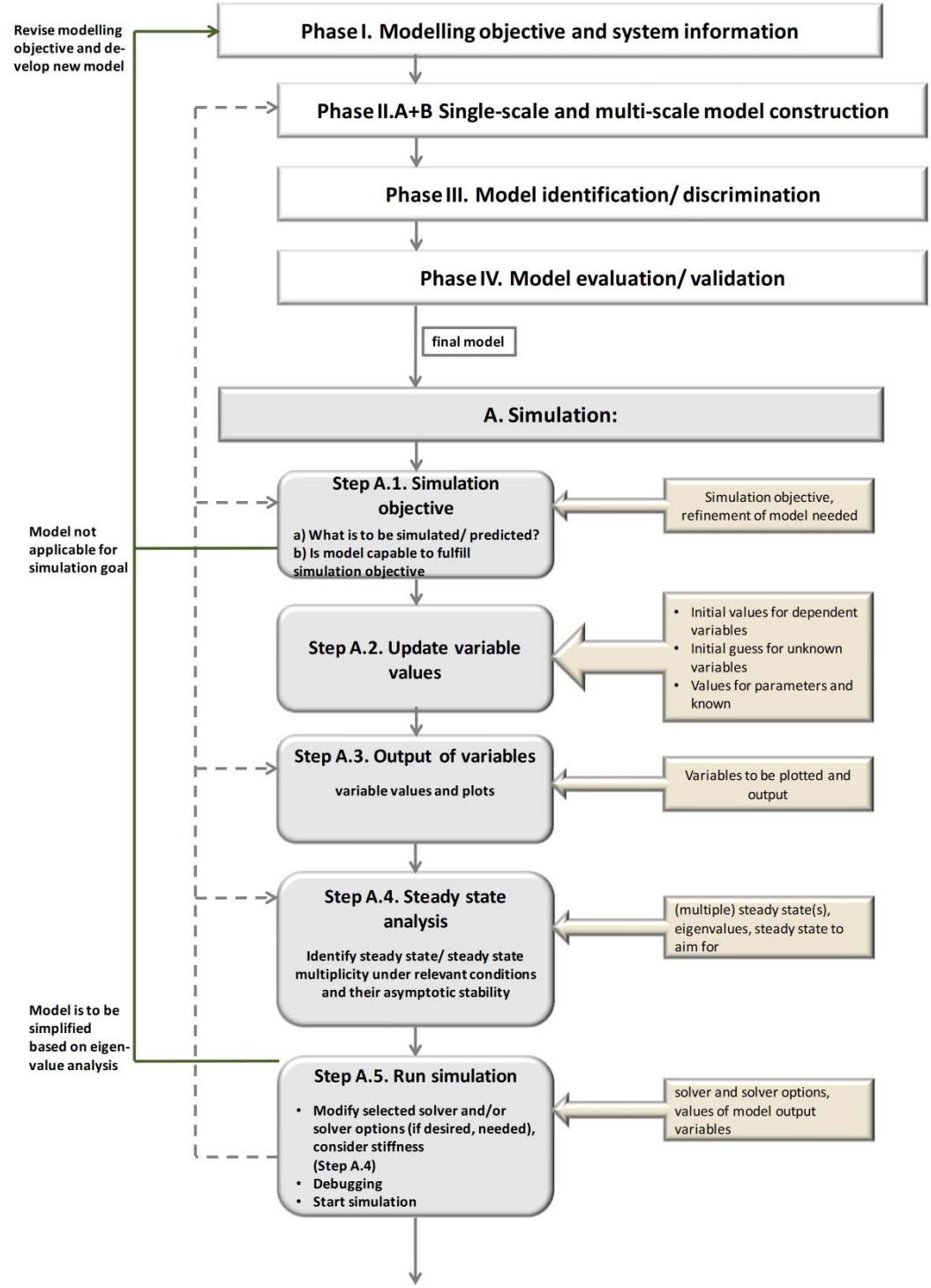


Figure 2.7 *Work-flow (boxes on left hand side) and data-flow (arrow boxes on right hand side) for model application for simulation (arrows on left indicate how Simulation is iterative).*

Step A.2. Update model variables

Objective: Adjust model variables to current model application.

Data: Values for parameters and known variables, initial values for dependent variables, initial guess for unknown variables.

Action-1: Update variable values (if applicable).

Action-2: Prepare documentation.

Step A.3. Output of variables

Objective: Set-up generation of output or plot of desired variables.

Action-1: Decide which variable values are to be written out during simulation.

Action-2: Decide which variables are to be plotted during simulation.

Action-3: For dependent output variables decide on output interval with respect to independent variable.

Action-4: Extend model by code needed for generation of specified outputs and plots.

Step A.4. Steady state analysis

Objective: Identify steady state (or steady state multiplicity) of model under relevant conditions and asymptotic stability of steady state(s).

Note: This step is only relevant if the stability and/or multiplicity of the system's steady state(s) is to be investigated for the relevant conditions.

Note: Both, the dynamic model or the corresponding steady state model can be used for this step.

Action-1: Update/modify solvers and solver options if applicable.

Action-2: Solve model for relevant conditions (if steady state model: initial guesses for unknown variables; if dynamic model initial values for dependent variables and, if applicable, initial guesses for unknown variables) and identify corresponding steady state.

> *Condition-2:* Steady state multiplicity is to be investigated.
>
> If *Condition-2* do: Repeat *Action-2* for different relevant conditions.

Action-3: Calculate eigenvalues of the Jacobian matrix of (dynamic or steady state) model for initial conditions/guesses and steady state solution(s) found.

Note: Although the eigenvalue analysis and potential model simplification have been already investigated during the model development process it is recommendable to repeat them at this point because the model performance can change if the model is applied for different conditions (e.g. different system, different kinetic rates, etc.).

Rule-3a: For eigenvalue analysis of an overall multi-scale scenario combine equations from all coupled scales.

Note: The eigenvalues of the Jacobian matrices of the separate scale models will not be the same as the eigenvalues of the Jacobian matrix for the overall system because the coupling elements of the scales are missing in the separated Jacobian matrices (see fluidized bed reactor case study, Section 5.3).

Rule-3b: Steady state is asymptotically stable if and only if real parts of all eigenvalues of Jacobian matrix are negative.

Note: An asymptotically unstable steady state is not stable if exposed to small perturbations.

Rule-3c: System contains oscillations if eigenvalues of Jacobian matrix have imaginary parts.

Rule-3d: If $\frac{|\lambda_{max}|}{|\lambda_{min}|} > 1000$ (the ratio of the maximum absolute real part and the minimum absolute real part of the eigenvalues) the system is regarded as 'stiff'.

Rule-3e: If system is stiff dynamic solver that can handle stiffness is required for later solution of dynamic system (Step A.5).

Rule-3f: Stiff systems indicate that the number of dominant eigenvalues is less than the number of equations. This indicates a potential for model reduction/simplification (see Phase II). If system is to be simplified go back to model development process (Phases I-IV).

Action-4: Decide which steady state is to be reached and set initial values of dependent variables (or for steady state model: initial guesses for unknown variables) accordingly.

Action-5: Prepare documentation.

Step A.5. Run simulation

Objective: Predict system behaviour under given conditions (model application)

Action-1: Update/modify solvers and solver options (identified during model development process) if applicable.

Action-2: Debug set-up of simulation problem and model.

Note: Possible error sources are (initial) values assigned to model variables (Step A.2), implementation of model outputs (Step A.3) and modified solvers and solver options (Step A.5). The model implementation itself has already been verified and debugged during the model development process (Phases I-IV). However, during model application problems might be detected that have been overlooked before.

Action-3: Start simulation.

Action-4: Evaluate results with respect to simulation objective.

Action-5: Prepare documentation.

2.2.2 B. Optimization

The model application work-flow for optimization has the goal to optimize the design of the actual system (e.g. unit operation, product, process) by fixing the parameters identified in the previous phase (Phase III) and varying the design variables (known variables). In case the design target cannot be met the process/system concept needs to be revised. Figure 2.8 depicts the work-flow to be followed if a developed model is to be applied for optimization.

The inputs for an optimization problem are the design objective, the design variables (known variables) to be optimized and their initial guesses, the objective function and the applied optimization algorithm.

The obtained outputs are the sensitivities of the objective function (and sub-terms of it) with respect to the design variables, the optimized values of the design variables, the value of the objective function and the results of the verification simulation (see Simulation work-flow) applying the optimized design variable values.

Step B.1. Design objective

Objective: Establish design objective.

Note: Possible design objectives can be to minimize the cost for a process or unit operation, to minimize energy consumption, to reduce the concentration of a certain compound below a boundary value, to maximize the yield within a reactor/system, etc.

Action-1: Formulate design objective.

> *Condition-1:* Available model not applicable for specific optimization problem.
>
> **Note:** Model might not have been specifically developed for this optimization problem.
>
> If *Condition-1* do: Go back to model development process, adjust modelling objective (Phase I, Step I.1) and based on that go through Phases I-IV to develop appropriate model.

Action-2: Choose design variables from list of known variables.

Action-3: Prepare documentation.

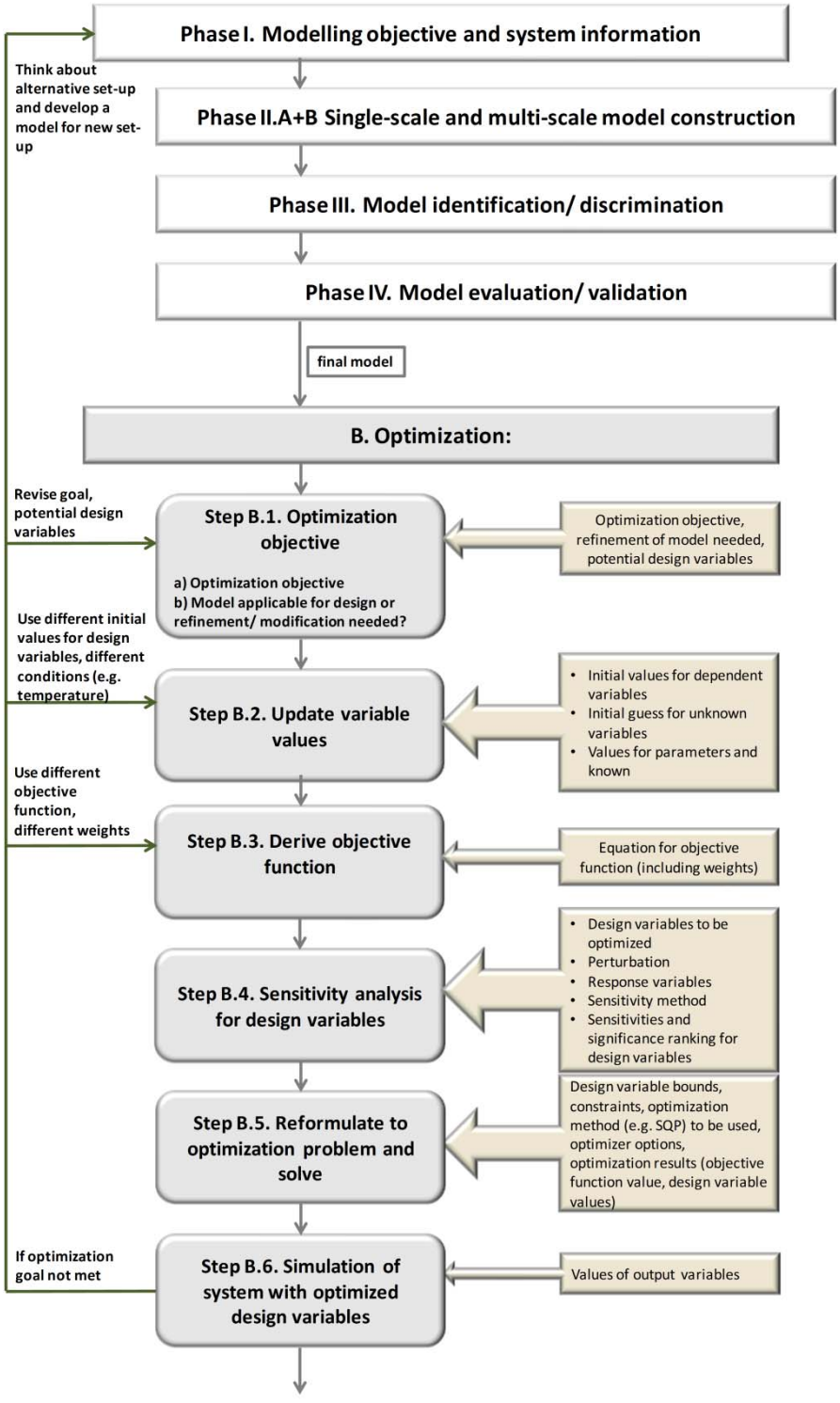


Figure 2.8 *Work-flow (boxes on left hand side) and data-flow (arrow boxes on right hand side) for model application for optimization (arrows on left indicate how Optimization is iterative).*

Step B.2. Update variable values

Objective: Adjustment of model variables to current problem.

Data: Values for parameter and known variables, initial guess for unknown variables, initial values for dependent variables.

Action-1: Update variable values.

Action-2: Provide initial guesses for design variables.

Action-3: Prepare documentation.

Step B.3. Objective function

Objective: Establish appropriate equation for objective function of design problem.

Action-1: Design appropriate objective function including possible weight factors that reflects design objective (see Step B.1).

Action-2: Extend model by equation for objective function.

Action-3: Prepare documentation.

Step B.4. Sensitivity analysis of design variables

Objective: Identification of design variables that have low/no impact on objective function ->decrease size of optimization problem and gain better understanding of system.

Note: Sensitivity analysis is an especially attractive tool for design problems with large numbers of design variables since it has the potential to reduce the complexity.

Note: Sensitivity analysis is not an obligatory step for optimization. If sensitivity analysis is omitted it is not possible to reduce the size of the optimization problem.

Action-1: Decide between local or global sensitivity analysis and for method.

Action-2: Connect implementation of chosen method with model.

> *Condition-2:* Software environment does not support sensitivity analysis or connection to external software/implementation of sensitivity analysis.
>
> If *Condition-2* do: a) change software environment, b) perform simulations needed for sensitivity analysis 1-by-1 manually (time-consuming), or c) omit sensitivity analysis for parameter estimation.

Action-3: Set design variables (see Step B.1) to perturbation variables for sensitivity analysis.

Action-4: Give boundaries for perturbation of design variables.

Action-5: Adjust solver selection and solver (e.g. output times of dependent variable values).

Action-6: Assign objective function (see Step B.2) as response variable for sensitivity analysis.

Action-7: Optional: assign additional variables related to the design objective as response variables (e.g. variables occurring in equation for objective function).

Note: Assigning additional response variables (next to objective function) for sensitivity analysis is for example of interest in a multi-objective optimization where different objectives are combined in one objective function in order to determine the sensitivity of the different objectives with respect to the design variables. In some cases a modification of the same design variable has a favourable impact on one objective but a negative impact with respect to a second objective (see combustion case study, Section 5.1).

Action-8: Perform sensitivity calculations.

Action-9: Evaluate obtained design variable significance ranking.

> *Condition-9:* Design variables with low sensitivity (compared to other design variables).
>
> If *Condition-9* do: Remove design variables with low sensitivity from list of design variables.

Action-10: Prepare documentation.

Step B.5. Reformulate to optimization problem and solve

Objective: Find optimal values of design variables.

Action-1: Provide/update initial values for design variables.

Action-2: Provide boundaries for design variables.

Action-3: Put constraints on other model variables if applicable.

Action-4: Decide for optimization algorithm (e.g. SQP).

Action-5: Create optimization problem with goal to minimize/maximize objective function (see Step B.3) subject to equality (model equations) and inequality constraints (design variable bounds, constraints on other model variables):

- Connect implementation of optimizer based on selected optimization algorithm to model and its numerical solver;
- Transform model equations, design variable bounds and specified constraints on other model variables to constraints of optimization problem;
- Implement and connect objective function calculation to optimization problem.

Action-6: Set up optimizer options (e.g. convergence tolerance).

Action-7: Set up or implement desired outputs (e.g. variable values and/or plots, at least: estimated design variable values and value of objective function).

Action-8: Perform optimization.

Action-9: Prepare documentation.

Step B.6. Simulation of system with optimized design variables

Objective: Verify that system performance meets design objective using optimized values of design variables.

Action-1: Go back to original model (without connected optimization problem).

Action-2: Set values of design variables to optimized values.

Action-3: Perform simulation (follow relevant steps of simulation work-flow, Section 2.2.1).

Action-4: Use simulation results to check if design objective is met by system.

Condition-4: Design objective not met.

If *Condition-4* do:

1. Go back to Step B.3 and use an alternative objective function.
2. Go back to Step B.2 and use different initial values for the design variables or use different values for the remaining variables. The latter represents a change of condition in the system (for example temperature, reactor length).
3. Go back to Step B.1 and adjust the optimization objective or add additional known variables to the list of potential design variables (e.g. reactor volume, temperature, etc.).
4. Invent a new system which might be better suited to meet the optimization goal and go back to the model development process (Phases I-IV) to develop a model for the new system. For example, an alternative reactor type or unit operation could be used

Action-5: Prepare documentation

Chapter 3. Computer-aided modelling framework

A computer-aided modelling framework has been developed based on the generic modelling methodology presented in Chapter 2. The overall objective of the computer-aided modelling framework is to increase the efficiency of the modeller with respect to time, resources, model-reuse and model quality. That is, to implement the generic modelling work-flows into a computer-aided software by identifying the required features and support a modelling tool needs to provide for each step of the work-flows. Furthermore, for each work-flow step the automation potential needs to be considered, that is, which steps or sub-steps can be automated by a modelling tool and completely performed by the computer.

In this chapter the computer-aided framework for the work-flows of the different phases of the modelling process (Phases I-IV) and model application for simulation and optimization is presented (see Figures 3.1-3.6). For each work-flow step the required features and the automation potential of a computer-aided modelling framework are highlighted. Figures 3.1-3.6 show the required features and support a modelling tool needs to combine for each work-flow step (in the brackets on the right hand side). In addition, for each step it is indicated how the modeller and the computer interact and which parts of the work-flow can be automated by a modelling tool.

The following features could support the modeller in the different steps of the work-flows:

- Incorporation of work-flow with work-flows for other tasks of iterative modelling process
- explanations (provided when needed) of the different work-flow steps and their available methods;
- automated generation of reports containing the provided information and obtained results from the steps performed;

Other additional features are indicated on specific work-flows.

3.1 Model development

3.1.1 Phase I. Modelling objective and system information

Figure 3.1 shows the 'Modelling objective and system information'-work-flow.

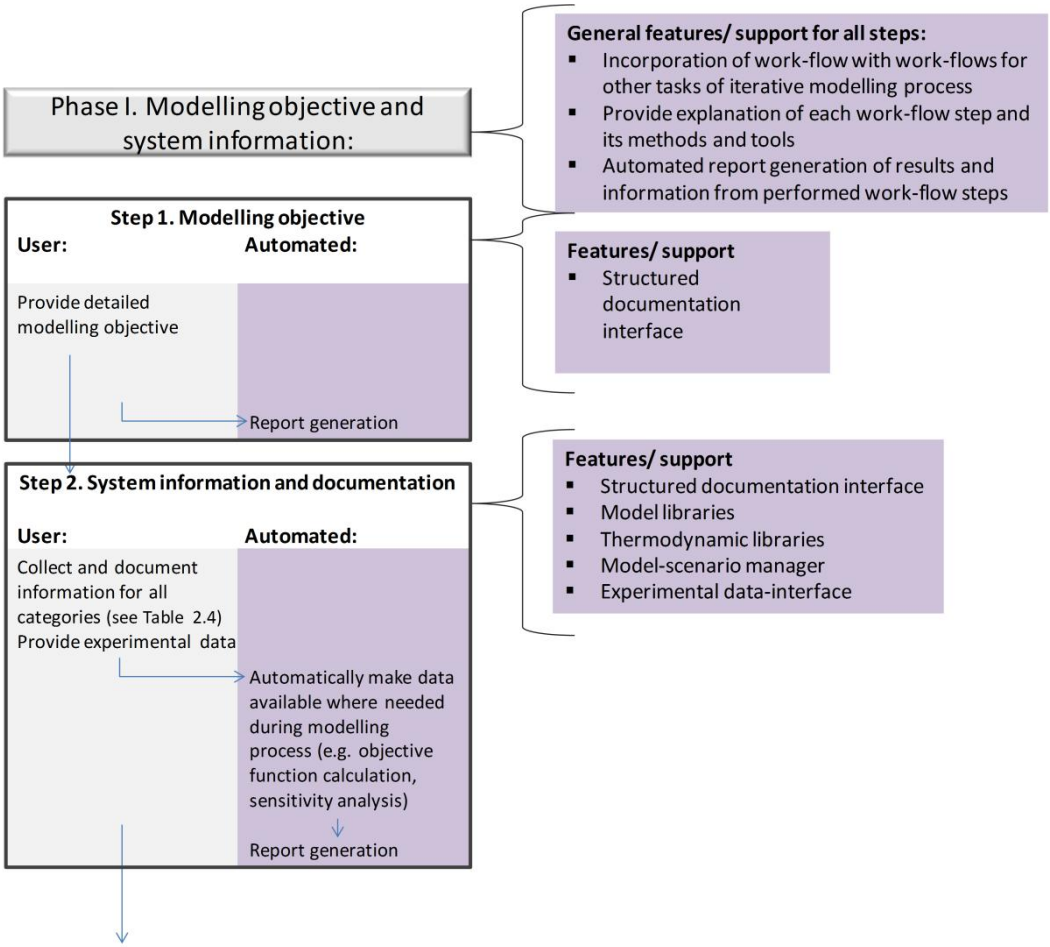


Figure 3.1 *Required features and automation potential identified for the steps of the work-flow for Phase I: Modelling objective and system information.*

As shown in Figure 3.1, a tool (interface) is needed for structured documentation to enter the model objectives. This part is manually but an interface to guide the modeller could be provided. An automatic report generator creates a document with the model objectives given by the modeller. For Step 2, three additional tools (model library, thermodynamic library, model-scenario manager) are needed. As in Step 1, the modeller needs to enter the data but the interface for documentation creates automatically the report, which is then combined with the

Step 1 report. The function of the library is the addition of apriori developed models and corresponding information. The function of the scenario manager is to manage different model scenarios of the same system.

3.1.2 Phase II. Model construction

II.A Single-scale model construction

Figure 3.2 shows the single-scale model construction work-flow including the potential for automation and the required features a computer-aided modelling framework ideally needs to provide. As shown in Figure 3.2, this work-flow needs additional features and tools, for example, a warning system in case required information from previous work-flow steps is missing or any other problem is detected.

The support a computer-aided modelling framework could provide for Step 1 is similar to the documentation interface described for Phase I (see Chapter 3.1.1). The modeller needs to provide the model description but the computer-aided modelling framework guides the modeller and automatically adds the provided information to the report.

For the derivation of the model equations (Step 2), a computer-aided modelling framework needs to incorporate a model library and features for systematic and automated generation of model equations based on modeller specifications (Jensen, 1998). In general Step 2 is manually. However, the model library provides a selection of apriori developed models from which the required model, a similar model or sub-models can be derived. In addition, an equation generation tool in many cases is able to automatically create the model equations based on the model description provided by the modeller in Step 1.

Features required for the model translation step (Step 3) are equation entering in a simple txt-syntax, tools for model translation (e.g. reverse polish notation (RPN)), discretization of PDEs, model aggregation, and model decomposition. Modeller inputs needed are specifications regarding discretization (e.g. method, number of discretization elements) and aggregation (e.g. which models are aggregated, variable communication between models) of models. Apart from these inputs, the model translation can be fully automated. The function of the model translation is the creation of model object that is readable by the computer from the txt-based model equations.

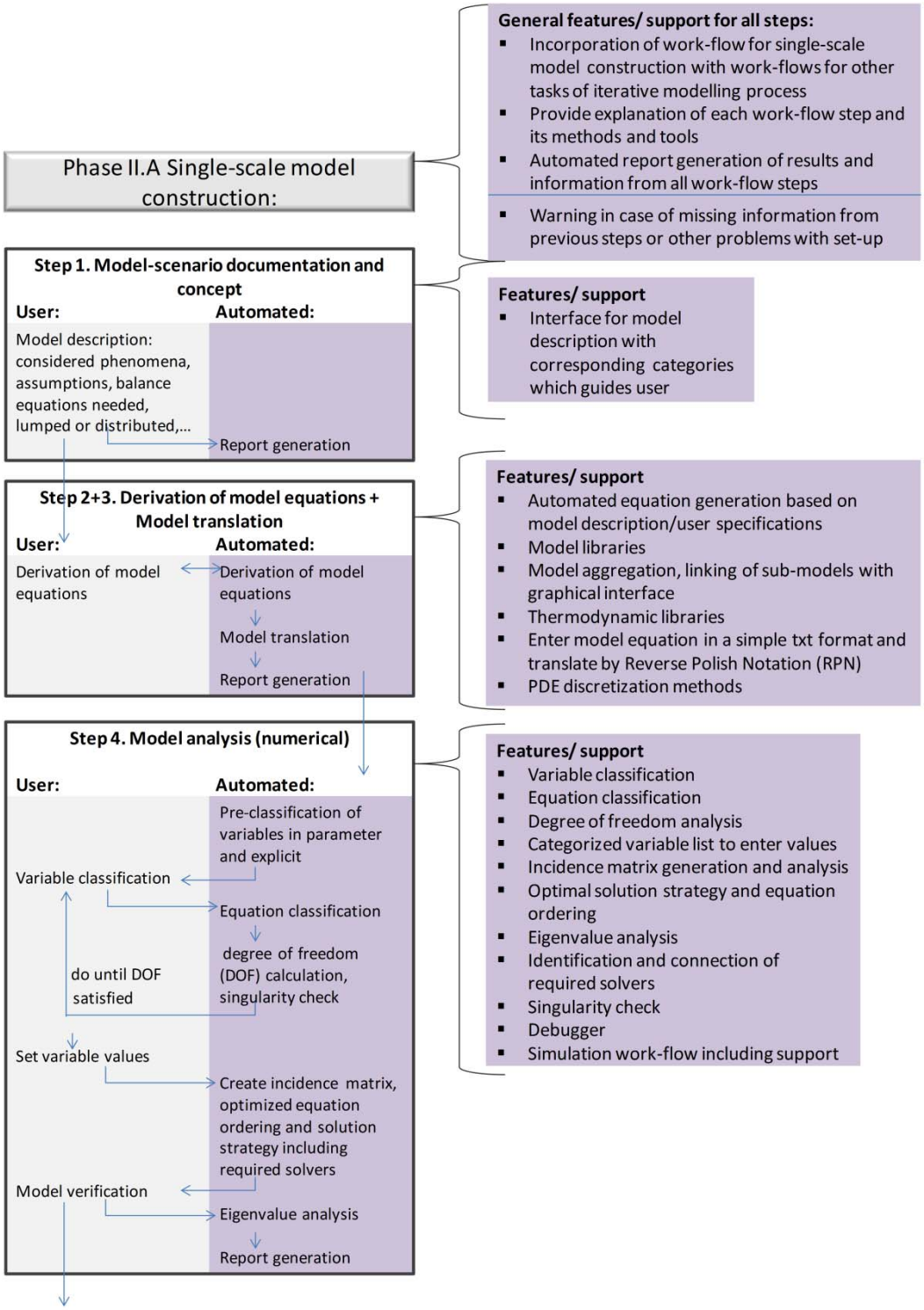


Figure 3.2 *Required features and support as well as automation potential identified for the steps of the work-flow for Phase II.A: Single-scale model construction.*

There are a large number of tools and methods a modelling framework needs to incorporate in order to support the modeller during the model analysis step (Step 4). Among these are tools for classification of variables and equations, for singularity check and, analysis of the degree of freedom (DOF), the display of the corresponding incidence matrix, optimization of the equation ordering, derivation of the solution strategy, solver selection, eigenvalue analysis, identification of opportunities for model decomposition and a debugger. As shown in Figure 3.2, the computer-aided modelling framework could automatically generate a pre-classification of the variables in non-explicit and explicit variables. But the modeller needs to specify which of the non-explicit variables are known, parameter, unknown or dependent and satisfy the degree of freedom (DOF). The calculation of the DOF as well as the derivation of the incidence matrix and a singularity check can be performed automatically by the computer-aided modelling framework for the current variable specification. Furthermore, the modelling framework needs to derive automatically the optimized equation ordering and the solution strategy including required solvers based on the variable classification and the incidence matrix. The modeller provides the required variable values manually and performs the debugging of the model.

II.B Multi-scale model construction

The identified features together with automation options that a computer-aided modelling framework should incorporate for the multi-scale model creation are summarized in Figure 3.3.
Support that a computer-aided modelling tool can provide for Step 1 (Scenario documentation and concept) is a structured documentation interface, a library of multi-scale modelling-scenarios, a data-flow scheme interface, connection to property prediction tools and thermodynamic libraries as well as the option to copy existing multi-scale scenarios. Using these supporting tools it is the modeller who has to provide and document the information while the computer-aided framework offers guidance and adds the provided information to the automatically generated report (like in Phase I, Phase II.A, Step 1).
Features needed for Step 2 are a model library and a data-flow interface. From this interface the modeller can connect models for the different scales and open the connected model files in order to develop the models following the work-flows of Phases I-IV.
The data-flow interface is also needed in Step 3. In addition to the features required for Step 2 the data-flow interface needs to allow analysis and modification of the data-flow between the scales as well as the addition of new scales. Another helpful feature for Step 3 is the multi-scale analysis tool that shows a list of the parameters and known variables for the current scale. In order to establish the data-flow the modeller needs to manually select the linked variables between the different scales from a list. For array variables the modeller needs to decide if the entire array is linked to the other scale or if the array variable is linked element-wise for multiple

calls of the other scale model. An example for the latter is a population balance model of an aerosol considering evaporation of the droplets. The evaporation sub-model needs to be called several times, once for each discrete droplet diameter, from the main model. The linked variables are arrays in the main-model and scalars in the sub-model. For each call of the sub-model the corresponding element of the input variable vectors from the main-model is communicated to the sub-model. If the modeller decides to add a new scale based on the multi-scale analysis the computer-aided modelling framework should automatically add the new scale and the linked variable(s) to the data-flow scheme.

Features a computer-aided modelling framework could provide for Step 4 are a graphical linking scheme interface, linking of models according to the linking scheme for different time and/or length scales, export of a developed model to a process simulator and eigenvalue analysis to help identifying different time-scales. The derivation of the linking scheme can be partially automated. The computer-aided modelling framework identifies the types of equations for each scale and the corresponding independent variables and based on this suggests a linking option (see Chapter 2.2, Step II.4). Furthermore, the computer-aided modelling framework needs to display the current linking scheme together with the linked variables for the different scales. The modeller selects the linking options for the linked variables (e.g. communicate entire array for each call of scale-model, communicate array element-wise for each call of scale model, form average value of an array before communication, multiply linked variables by a certain factor) and scales (e.g. call scale-model several times). The computer-aided modelling framework automatically links the model according to the specifications made, adjusts the array sizes in the scale-models and adds the required linking equations.

For Step 5, the features and automation that a computer-aided modelling framework can provide is the same provided for single-scale model development (see Step 4 in Figure 3.2).

Templates superimposed to the work-flow steps for different multi-scale problems/systems (for detailed example see fragrance aerosol case study, Section 5.4) are useful for all work-flow steps.

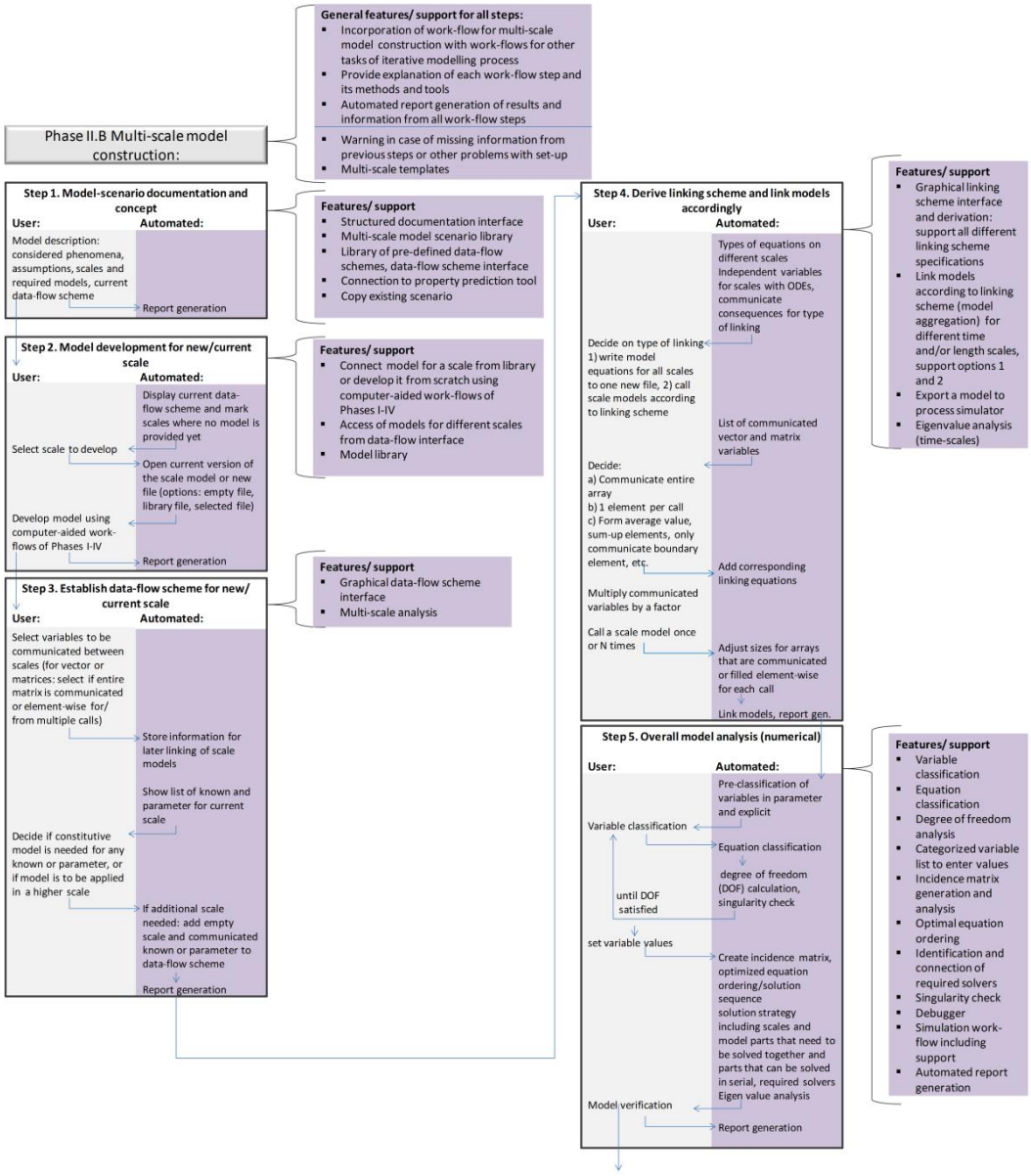


Figure 3.3 *Required features and support as well as automation potential identified for the steps of the work-flow for Phase II.B: Multi-scale model construction.*

The developed modelling framework offers support for all three key tasks for multi-scale modelling proposed by Ingram et al. (2004):

Task 1: Identification of the time and length scales to be considered;

Task 2: Retrieval or development of the sub-models for the considered scales;

Task 3: Linking of the sub-models to form an overall multi-scale model.

An overview is given in Table 3.1.

Table 3.1 *Support of computer-aided modelling framework for key tasks (Ingram et al., 2004) in multi-scale modelling*

Support of framework for Task 1. (Identification of time and length scales to be considered):	***Support of framework for Task 2. (retrieval or development of the models for the considered scales):***	***Support of framework for Task 3 (linking of the sub-models to form an overall multi-scale model):***
Computer-aided work-flow structure that efficiently and systematically guides modeller through the iterative process of derivation, construction, identification, discrimination, evaluation and comparison of alternative multi-scale scenarios starting from low complexity until the performance is satisfactory with respect to the modelling objective.	Computer-aided work-flow structure that systematically guides the modeller through model documentation, construction, identification, evaluation and validation	Data-flow and linking scheme interface which supports all linking scheme specification
System information in Phase I: sketch, systematic collection of possible phenomena and assumptions, collection of information on how system has been modelled before and provision of degree-of-detail-determining factors to support modeller in establishing starting scenarios to develop and iteratively improve with respect to the modelling objective	Model library Equation generation Linking of sub-models Simple syntax to enter model equations as close as possible to writing of equations in scientific papers, model translation by RPN, etc.	Linking of models according to established data-flow and linking scheme
Eigenvalue analysis to identify different time-scales	Numerical model analysis, debugging	Numerical model analysis of overall multi-scale scenario
Uncertainty analysis to reveal prediction quality of the models		
Sensitivity analysis to support finding of bottlenecks in the model and to add or remove scales		
Multi-scale analysis		
Multi-scale scenario library		
Multi-scale templates for specific problems/ systems		

3.1.3 Phase III. Model identification/ discrimination

Figure 3.4 gives an overview over the computer-aided work-flow for model identification.

For Step 1 a computer-aided modelling framework should provide an interface for experimental data which can handle different forms of data as well as computer-aided work-flows for data analysis and design of experiments for parameter estimation and model discrimination. The modeller needs to input (newly) available experimental data while the computer-aided modelling framework allows connection of the measurements to the corresponding model variables and automatically uses or provides the data wherever needed during the modelling process. The provided data is included in the generated reports.

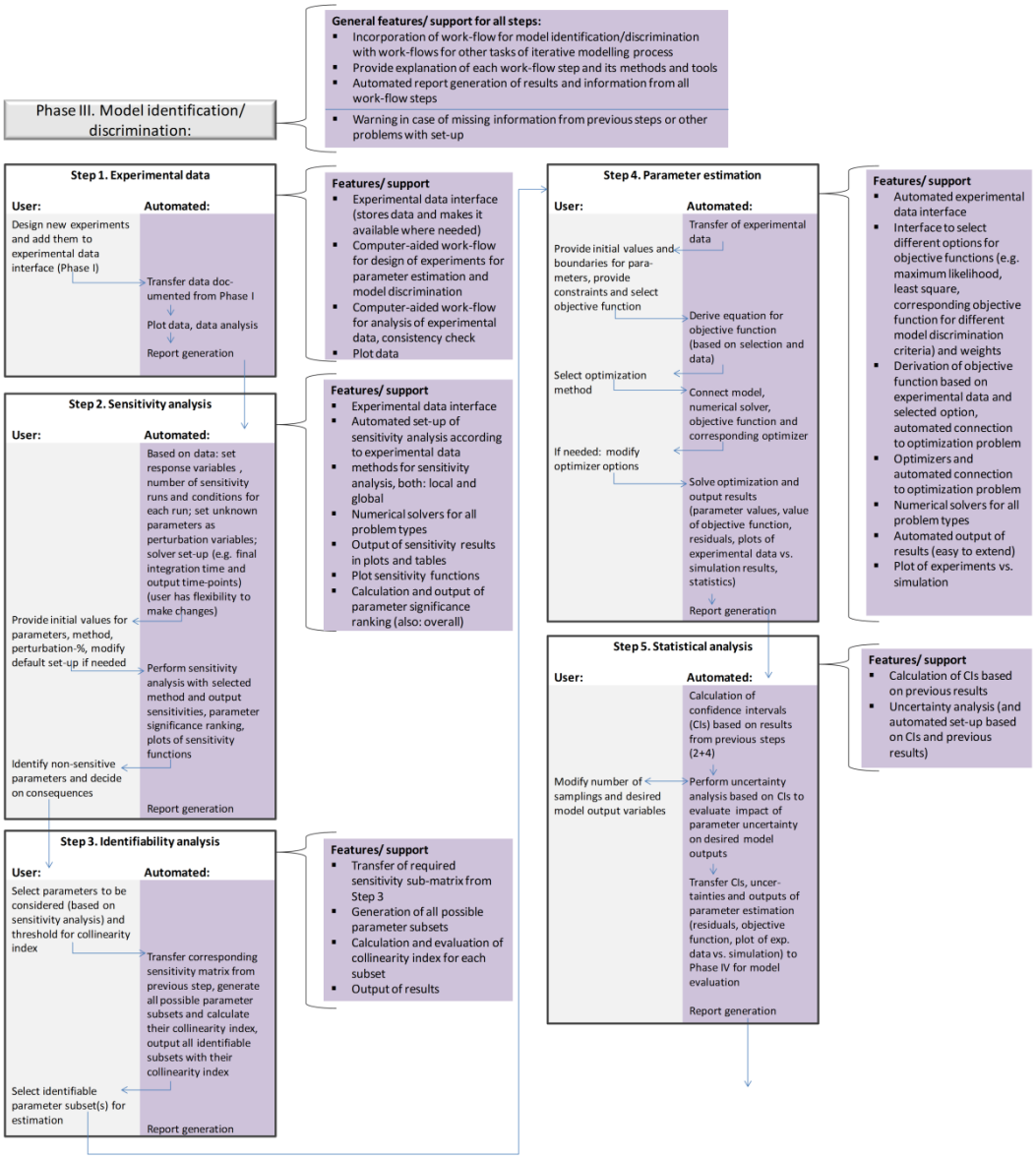


Figure 3.4 *Required features and support as well as automation potential identified for the steps of the work-flow for Phase III: Model identification/discrimination.*

Additional features required for Step 2 are tools for local and global sensitivity analysis, numerical solvers for all problem types and output of the sensitivity analysis results (e.g. plots of sensitivity functions, parameter significance ranking). Except for initial values of the unknown model parameters, the sensitivity analysis can be set-up and performed fully automated by the computer-aided modelling framework based on the variable specification and the entered experimental data. It is however important to still allow the modeller to flexibly adjust the automated set-up of the sensitivity analysis. Furthermore, the transfer and re-use of the

sensitivity analysis results where required in the following steps of the work-flow has to be automated. This, among others, holds true for the next step, the identifiability analysis.

For Step 3 the computer-aided modelling framework needs to provide features to generate all possible parameter subsets, calculate and evaluate their collinearity index (Brun et al., 2002) and output of the results. Also this step can be performed automatically by the computer-aided modelling framework.

For Step 4 the main additional features needed are an interface for the selection of the objective function, connection of optimizers, output of results (including statistics).

The objective function selection interface should allow the modeller to select the objective function type including weights and error model. Furthermore, the modeller needs to select the optimization method (e.g. SQP) and provide initial values and boundaries for the unknown model parameters as well as possible constraints for other model variables. Based on these specifications, the computer-aided modelling framework needs to automatically derive the equation for the objective function and set-up the actual parameter estimation problem by connecting the model and its numerical solver with the user-selected objective function and the optimizer corresponding to the selected optimization method. After the optimization the results are written out automatically, which are, the estimated parameter values, the value of the objective function, the residuals and a plot of the experimental data compared to the simulation results.

To support the modeller in Step 5 a modelling tool needs to provide statistical methods like confidence interval calculation and uncertainty analysis. Again, the set-up of the confidence interval calculation and the uncertainty analysis can be partly automated using the obtained results from the previous work-flow steps.

3.1.4 Phase IV. Model evaluation/ validation

For the model evaluation phase no detailed computer-aided work-flow has been developed and consequently, in this section only a list of useful features is given. Sensitivity analysis and uncertainty analysis are of importance. Furthermore, a modelling tool should include a number of different model discrimination criteria and should automatically determine and output these measures in case the corresponding objective function has been used during the parameter estimation (Phase III). A scenario manager which allows the storage, modification and access of different scenarios of the same problem is of great help. It should be possible to automatically generate an evaluation report for each model scenario and a comparison report for the different modelling scenarios considered.

3.2 Model application

3.2.1 A. Simulation

The main features a computer-aided modelling framework needs to provide to support the modeller during the simulation process are briefly summarized in Figure 3.5.

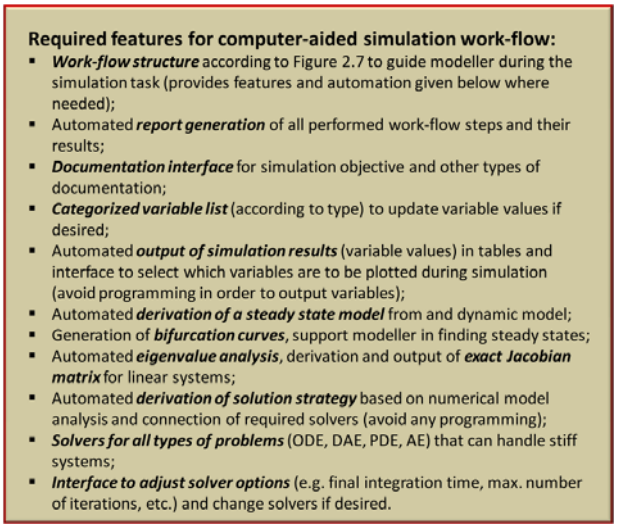


Figure 3.5 *Required features and support as well as automation potential identified for the steps of the work-flow for model application A: Simulation.*

3.2.2 B. Optimization

Figure 3.6 shows the work-flow for a modelling tool that can support the modeller during the solution of an optimization problem. This figure also indicates the features that a modelling tool should ideally include.

The required features and their automation potential have already been highlighted during the previous work-flows (see Phases I-IV). In addition, the optimization work-flow should allow the automated transformation of the model implementation into an optimization problem (Step 5) as well as the reverse transformation back to a simulation problem (Step 6).

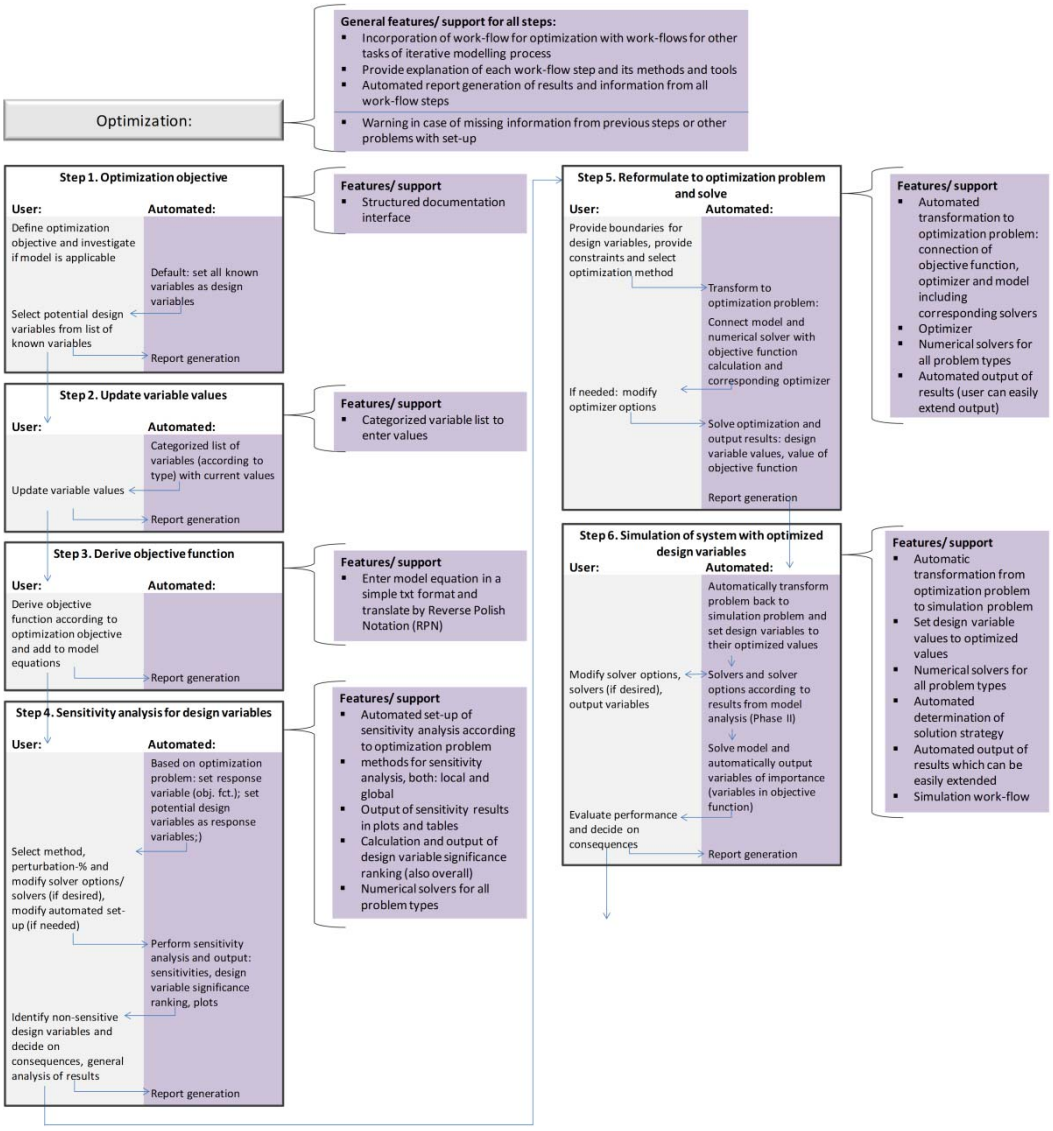


Figure 3.6 *Required features and support as well as automation potential identified for the steps of the work-flow for model application B: Optimization.*

3.3 Summary

The developed computer-aided modelling framework which is structured based on the above described computer-aided work-flows combines features spanning all three structural layers proposed by Von Wedel et al. (2002) as requirements for modern modelling tools (see Section 1.2.2). In Table 3.2 the most important features of modelling tools that have been identified for the different work-flows are classified into these three layers.

Table 3.2 *Categorization of most important modelling tool features identified for the different work-flows in model development and application based on their layer*

Mathematical base layer	***Systems engineering layer***	***Chemical engineering layer***
Computer-aided work-flow structure (especially for single-scale model construction, model identification and application)	Linking schemes, model aggregation, computer-aided work-flow for multi-scale modelling	Model documentation interface, Computer-aided work-flow interface (especially for Phase I 'Modelling objective and system information')
Explanation of work-flow steps and methods used	Incidence matrix generation	Equation generation feature
Introduction of model equations in simple txt-syntax and translation by RPN	Model scenario manager	Model libraries
PDE discretization methods		Thermodynamic libraries
Numerical model analysis features		Property prediction tool
Generic numerical solver, optimized equation ordering and derivation of solution strategy		Multi-scale scenario library
Optimizer		Multi-scale templates for specific problems
Automated transformation of a simulation problem into an optimization problem and vice versa		Export of models to process simulator
Sensitivity analysis		Automated report generation and documentation
Identifiability analysis		
Confidence interval calculation		
Uncertainty analysis		
Experimental data interface		

Chapter 4. Implemented modelling tool and software architecture

The developed computer-aided modelling framework has been implemented into an user-friendly software. The resulting modelling tool is called ICAS-MoT ('Modelling Testbed') and is integrated within the ICAS ('Integrated Computer-Aided System') software (Gani et al., 1997) which combines a number of tools, for example, property prediction, solvent design, process simulation, equation generation and thermodynamic databases (see Figure 4.1).

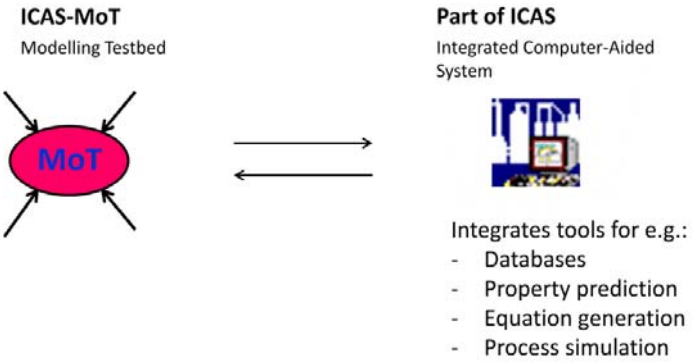


Figure 4.1 *ICAS-MoT ('Modelling Testbed').*

MoT had already been developed prior to the start of this Ph.D. project by Russel & Gani (2000), Sales-Cruz & Gani (2003) and Sales-Cruz (2006). However, in order to incorporate the developed modelling methodology (Chapter 2) and the corresponding computer-aided modelling framework (Chapter 3), the computer-aided work-flow structure and interface had to be implemented, additional automation and modelling features were added while existing features and methods have been improved. A summary of the modifications and extensions of MoT made within this Ph.D.-project is given in Appendix A.1.

This chapter is divided in two parts. Section 4.1 describes the software architecture and features of MoT using the computer-aided work-flows for single- and multi-scale model construction, model identification and model application for problems requiring optimization. Section 4.2

refers to features and ideas that have not yet been implemented in MoT mainly related to the multi-scale model construction work-flow.

4.1 ICAS-MoT

This section provides an overview of MoT and its main features with the goal to give an impression on how the developed computer-aided modelling framework has been turned into reality. Detailed information on MoT and all its features is given in the MoT manuals provided together with the ICAS software.

First the general structure of the modelling tool ICAS-MoT is described before going to the detailed computer-aided work-flows. Figure 4.2 gives an overview of the structure and the integration of MoT in ICAS.

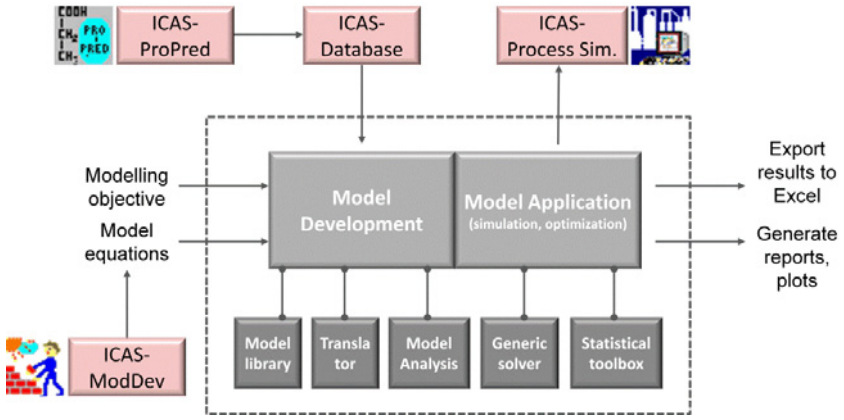


Figure 4.2 *Basic structure of MoT and its connection with ICAS.*

Like the developed modelling framework (Chapter 3), MoT consists of two main parts: Model development and model application. For these two parts different toolboxes, which work in the background of the work-flow interface, are needed. The main toolboxes for the model development process are the model library, model translator and model analysis. The model application part requires a generic solver and a statistical toolbox that for example provides tools for sensitivity and uncertainty analysis. In MoT the modelling objective together with the model equations need to be provided by the user. However, MoT is connected to a model library and the ICAS-tool ModDev (Jensen & Gani, 1996; Jensen, 1998) that can generate equations based on user specifications. In general it is straight forward to introduce a model in MoT because the equations are introduced through a very simple syntax that is as close as possible to the way equations are written in scientific papers. Therefore, no programming is necessary to develop, analyse and/or solve a model in MoT. Apart from ICAS-ModDev there are

also connections to the process simulator and the thermodynamic databases of ICAS. The results obtained during the model development and application process in MoT can be exported to Microsoft Excel. In addition MoT creates plots and reports.

The structure of the work-flow interface is given in Figure 4.3. It shows the work-flow selection window of MoT. This is the starting screen when MoT is launched.

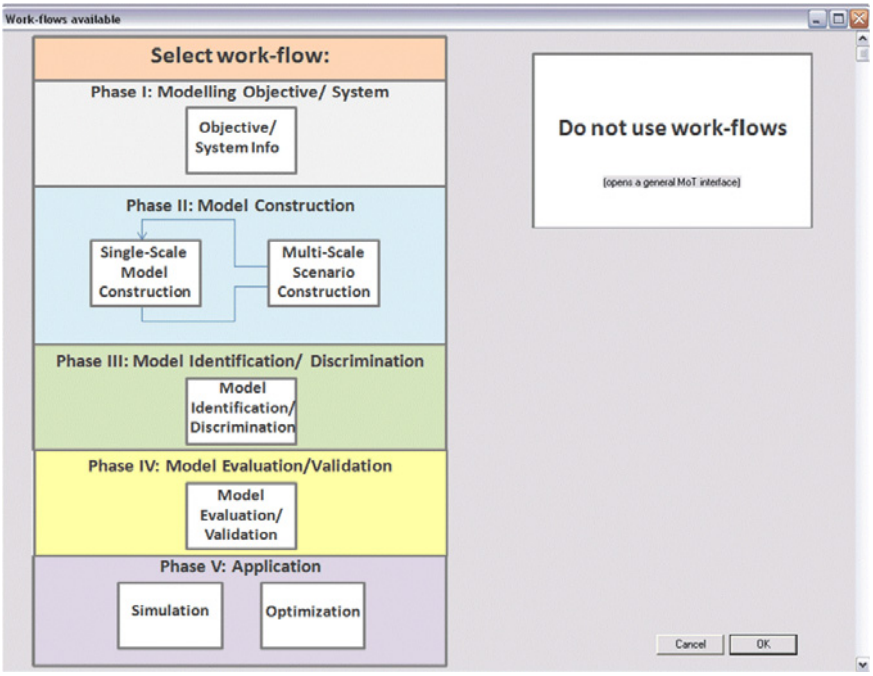


Figure 4.3 *Work-flow selection window in ICAS-MoT.*

The modeller selects the desired work-flow according to the current state of the model and the desired task to be performed. The main MoT-window opens and contains the selected work-flow (the left hand side of Figure 4.4) and the translated model equations (right hand side of Figure 4.4).

At each step of the work-flow MoT provides the required support identified in Chapter 3. MoT provides background information about the different work-flow steps, the related modelling features and the needed tools to the modeller (see Figure 4.5).

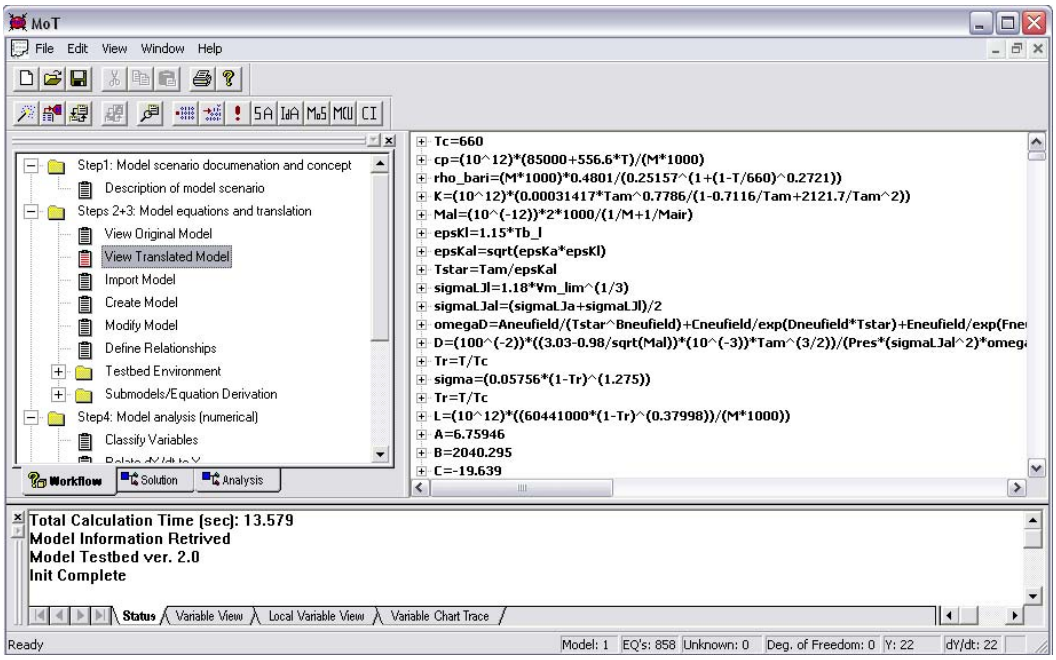


Figure 4.4 *Main MoT window with loaded work-flow for single-scale model construction.*

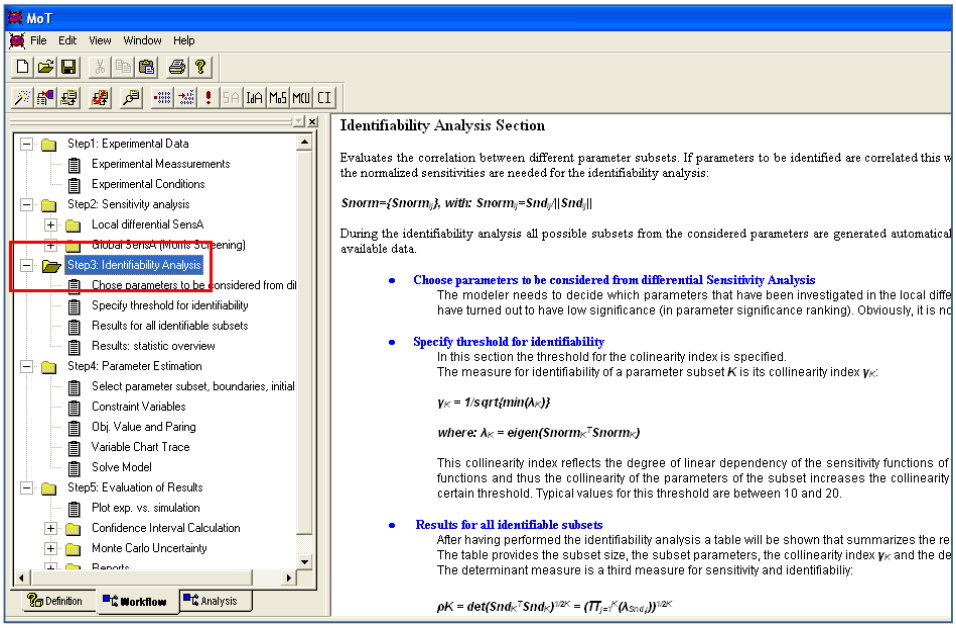


Figure 4.5 *Information given to modeller for different work-flow steps and corresponding methods.*

The work-flow manager (see Figure 4.6) allows the user to switch between the work-flows during model development and application.

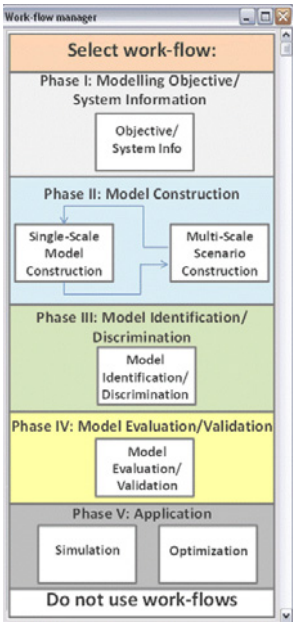


Figure 4.6 *Work-flow manager in ICAS-MoT (switching between work-flows).*

In addition to the work-flow based interface it is also possible to use the general MoT interface, which combines all MoT features but does not provide any work-flows and therefore less automation and support.

4.1.1 Phase I. Modelling objective and system information in MoT

For the 'Modelling objective and system information'-phase MoT provides a structured documentation interface (Figure 4.7). The information provided here is written to the automatically generated report. This phase has two steps. In Step 1 the modeller states the model objectives. In Step 2, the modeller collects and documents the system information. This includes the entering of steady state or dynamic experimental data and relating it to the corresponding model variables. The entered experimental data will be available wherever needed during the model development and application. Note however, that Step 2 is optional, that is, data is entered only if available.

In summary, the essential features a modelling tool should provide to support this first phase of the modelling process identified in Chapter 3 (see Figure 3.1) are made available in MoT.

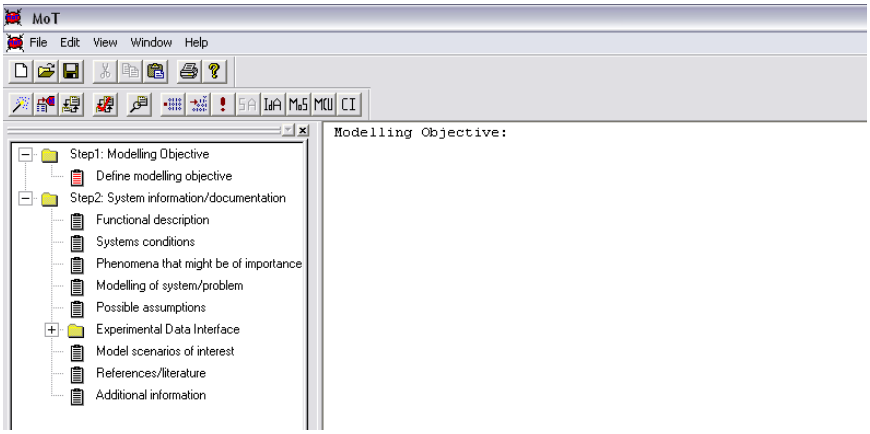


Figure 4.7 *Structured documentation interface (work-flow for Phase I).*

4.1.2 Phase II. Single-scale and multi-scale model construction in MoT

Single-scale model construction in MoT

The implementation of the computer-aided single-scale model construction work-flow (see Figure 3.2) in MoT is presented in detail here. It consists of four steps: 1. Model scenario documentation and concept, 2. Derivation of model equations, 3. Model translation, 4. Model analysis (numerical).

MoT provides the required features, guidance and automation identified for this workflow in Chapter 3 (see Figure 3.2). The main features and the work-flow interface are highlighted here. Figure 4.8 gives a screenshot that shows the single-scale model construction work-flow loaded to the MoT interface (left hand side) and a dialog window to support the provision of the model description needed for Step 1 of the single-scale model construction work-flow (right hand side). Figures 4.9 and 4.10 illustrate the interfaces for the model derivation and the model translation steps (Steps 2 and 3), respectively. The model equations are introduced to the equation editor in a simple text-syntax and are translated by MoT applying reverse polish notation (RPN). In order to support the modeller during the derivation of the model equations a link to open the equation generation tool ICAS-ModDev is available (see Figure 4.9). Furthermore, model libraries are also available.

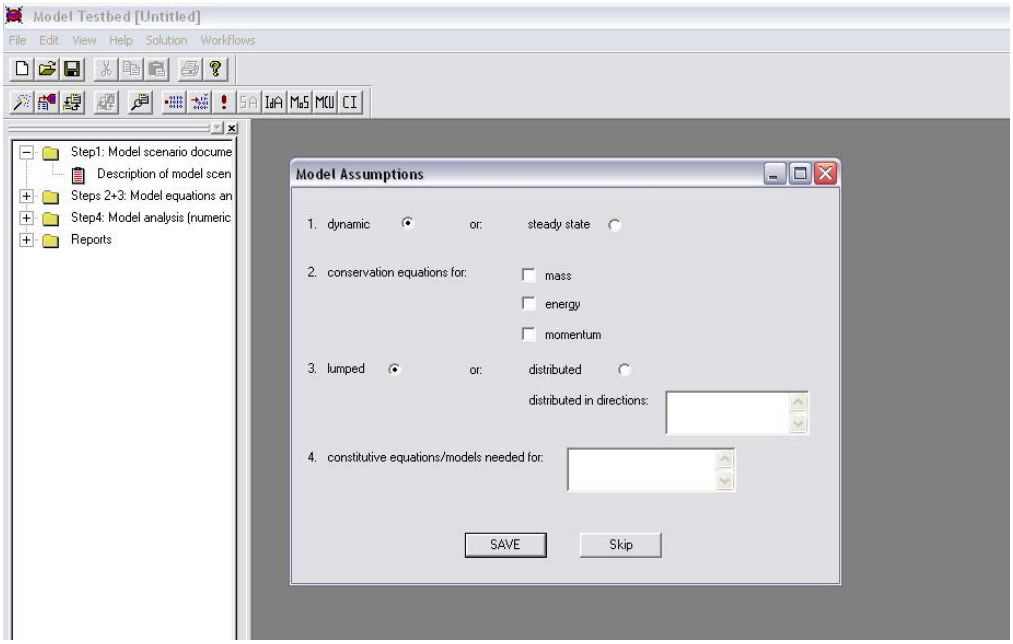


Figure 4.8 *Step 1. Model scenario documentation and concept.*

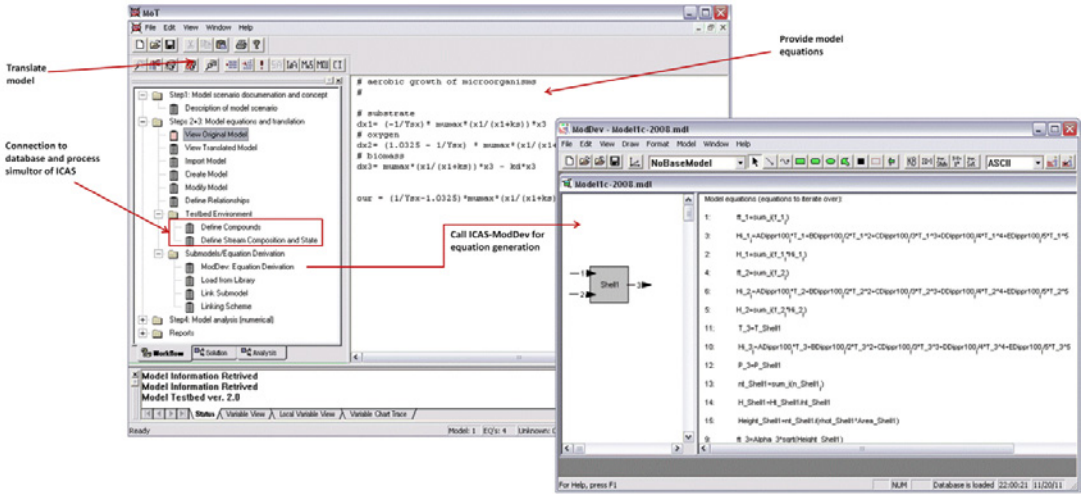


Figure 4.9 *Features to support derivation of model equations (Step 2).*

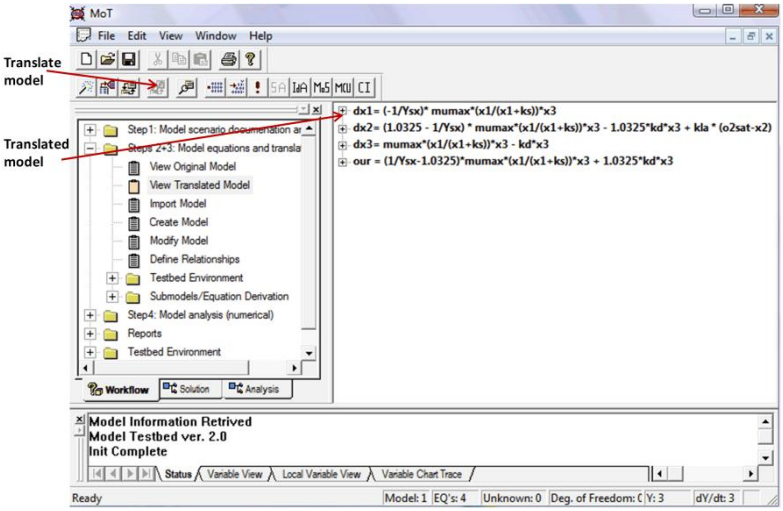


Figure 4.10 *Translated model in MoT (Step 3).*

The next step is the (numerical) model analysis (Step 4). The main screen for model analysis is shown in Figure 4.11. The list of explicit variables is created automatically during model translation. The modeller needs to classify the remaining model variables as either parameter, known, unknown or dependent. For each change of the variable classification the degree of freedom is updated and displayed (lower right corner of screenshot) and a singularity check is performed. Furthermore, the incidence matrix is displayed. The modeller has the option to ask MoT to re-organize the equations to an optimal order. A model development report of all performed steps and their results is generated automatically.

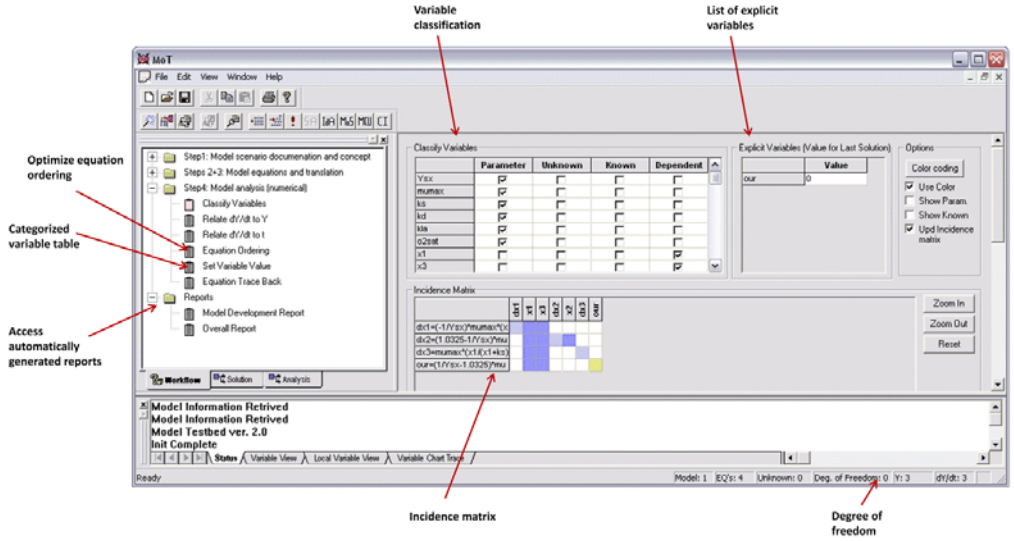


Figure 4.11 *Main interface for numerical model analysis (Step 4).*

The features and automation MoT provides for single-scale model development are also required by the multi-scale model construction work-flow. However, the multi-scale model construction work-flow has a different structure (see also Chapters 2 and 3) and requires some extra features and support.

Multi-scale model construction in MoT

Only the main (extra) features and support MoT provides for the multi-scale model construction work-flow are presented here.

Figure 4.12 shows a screenshot of the multi-scale modelling work-flow. Here, the two main features are: establish the data-flow and the linking scheme for the model. The linking scheme type currently supported by MoT connects a main-scale ('Scale 1' in Figure 4.12) with a sub-scale ('Scale 2' in Figure 4.12). The sub-scale model is called from the main-scale model. For the linked models, the transfer of data (data-flow) needs to be established.

Data-flow scheme

The objective of the data-flow scheme is to establish the linked variables between two scales. A detailed description how a data-flow scheme should be implemented together with an application example is given in Section 3.1.2 (Phase II.B). The data-flow scheme interface in MoT is shown in Figure 4.12 (top). It allows the connection of a MoT-file to a sub-scale. This can be an already developed MoT-model (e.g. from library) or, in case the model for the sub-scale still needs to be developed, an empty MoT-file can be opened. The arrow-fields in the data-flow scheme shown in Figure 4.12 are used to establish the data-flow between the two scales. The linked variables can be selected from a list. For array variables the modeller has the option to link the entire array to the sub-model or to call the sub-model multiple times and link one element of the array for each call. The array sizes are automatically adjusted to the number of calls specified (see linking scheme). If a linked variable does not exist in the main-model it is created. For all automated modifications MoT outputs a message to the modeller.

Linking scheme

The objective of the linking scheme is the establishment of the linking options between two scales (e.g. call sub-scale multiple times). A detailed description how a linking scheme should be implemented together with an application example is given in Section 3.1.2 (Phase II.B).

The linking scheme interface of MoT is shown in Figure 4.12 (bottom). It allows to specify and modify the number of calls of the sub-scale model (e.g. for population balance models with changing particle size distributions) from the main model. The default value is 1. If the number

of calls is modified the dimensions of all data-flow variables are adjusted automatically and the performed changes are communicated to the modeller.

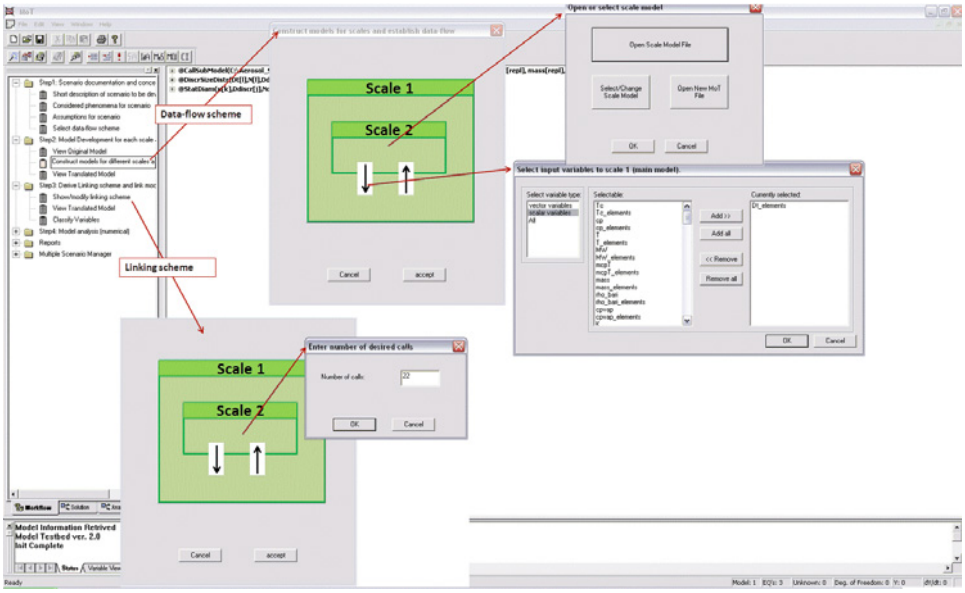


Figure 4.12 *Multi-scale work-flow in MoT: Establishment of data-flow scheme and linking scheme.*

Solver options interface

A solver options interface allows to specify the solver options for the call of the sub-scale model. In case of multiple calls of the sub-scale it is possible to provide different solver options for each call.

4.1.3 Phase III. Model identification/ discrimination in MoT

For the model identification work-flow all features and the automation strategy developed in Chapter 3 (see Figure 3.4) have been implemented in MoT, with the exception of design of experiments and data analysis. In the following, the implementation of the computer-aided model identification work-flow is highlighted for each work-flow step through screenshots of the most important features.

Step 1: Here, the data to be used for model identification/ discrimination is entered. The data is divided into two parts (Figure 4.13)

- measured (experimental variables);
- conditions (variables representing different experimental conditions).

The entered data is included in the generated report.

MoT provides a link to the experimental data interface. The modeller can access and update the data given during system information step in Phase I. Figure 4.13 shows a screenshot of the main MoT-window with the work-flow for model identification and the experimental data interface.

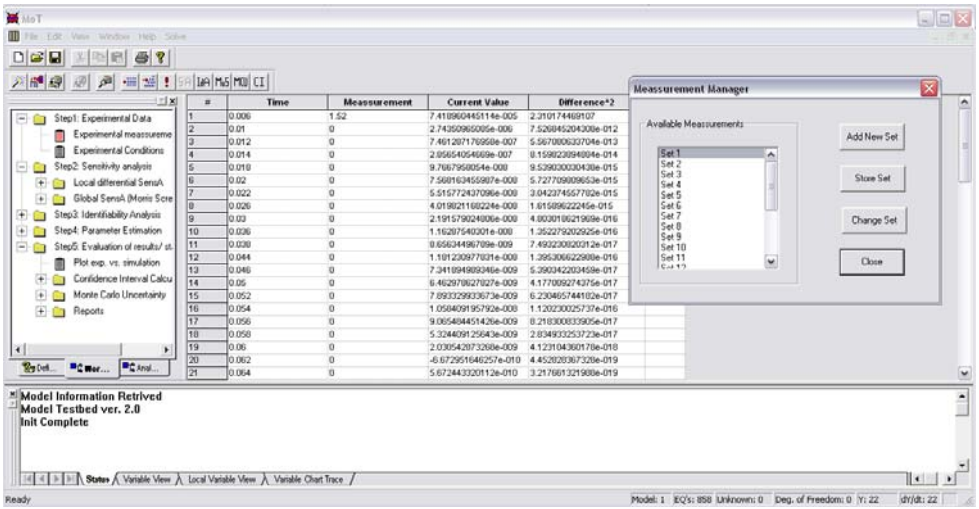


Figure 4.13 *Model identification work-flow and experimental data step.*

Step 2: Here, a sensitivity analysis is performed. This step is optional. Two alternative methods have been implemented: a local and a global method. The first option is a local differential sensitivity analysis Through this feature, MoT is able to automatically set-up the sensitivity analysis based on the experimental data with the option to modify any of the automated steps. Figures 4.14 and 4.15 illustrate the automatic selection of the response variables and the set-up of multiple sensitivity output-times, respectively according to the experimental data provided. For the local differential sensitivity analysis the sensitivity parameters are perturbed one by one forward and backward by the specified perturbation steps and the model is simulated for the corresponding perturbed parameter values. The resulting sensitivities for the different response variables are calculated at each specified sensitivity output point. The absolute and non-dimensional sensitivities Sa and S_{nd} are calculated according to Equations 4.1 and 4.2 (Brun et al., 2002; Sin & Vanrolleghem, 2007; Sin et al., 2010):

$$Sa(i,j) = \frac{Y_{Fi} - Y_{Bi}}{2 \cdot \Delta P_j}, \quad i = 1, NVAR; j = 1, NPAR \tag{4.1}$$

$$S_{nd}(i,j) = Sa(i,j) \cdot \frac{P_j}{Y_i}, \quad i = 1, NVAR; j = 1, NPAR \tag{4.2}$$

Here, Y_{Fi} and Y_{Bi} are the response variables, the indices F and B stand for forward and backward, respectively. Whereas the index i is the response variable counter and j is the parameter counter. ΔP_j is the absolute perturbation value of the parameter j. The non-

dimensional sensitivity S_{nd} (so called relative sensitivity function, Sin & Vanrolleghem, 2007) is normalized by the initial value of the parameter P_j and model output variable Y_i at the current data point and the initial value of the parameter.

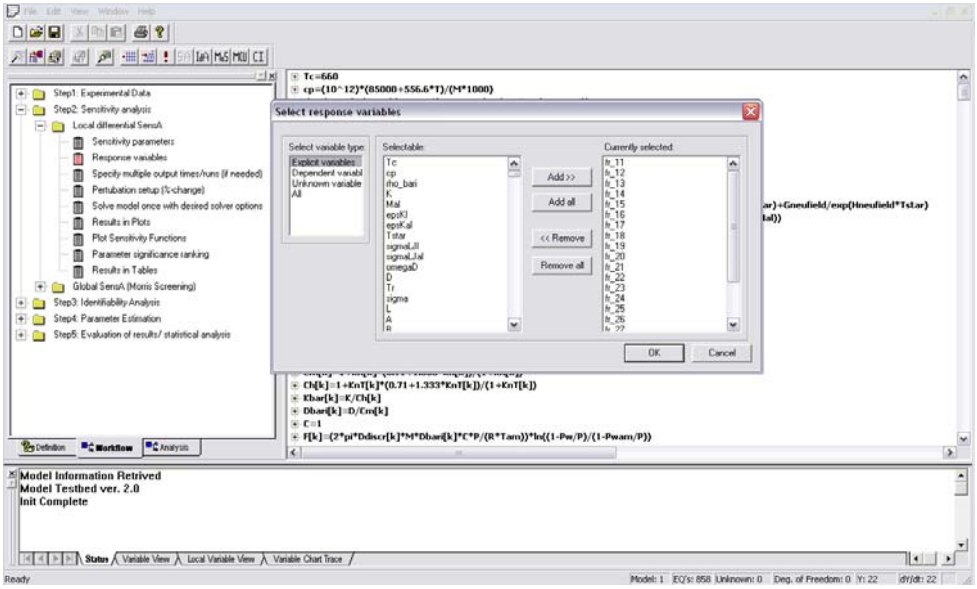


Figure 4.14 *Automated selection of response variables for sensitivity analysis according to experimental data introduced to MoT.*

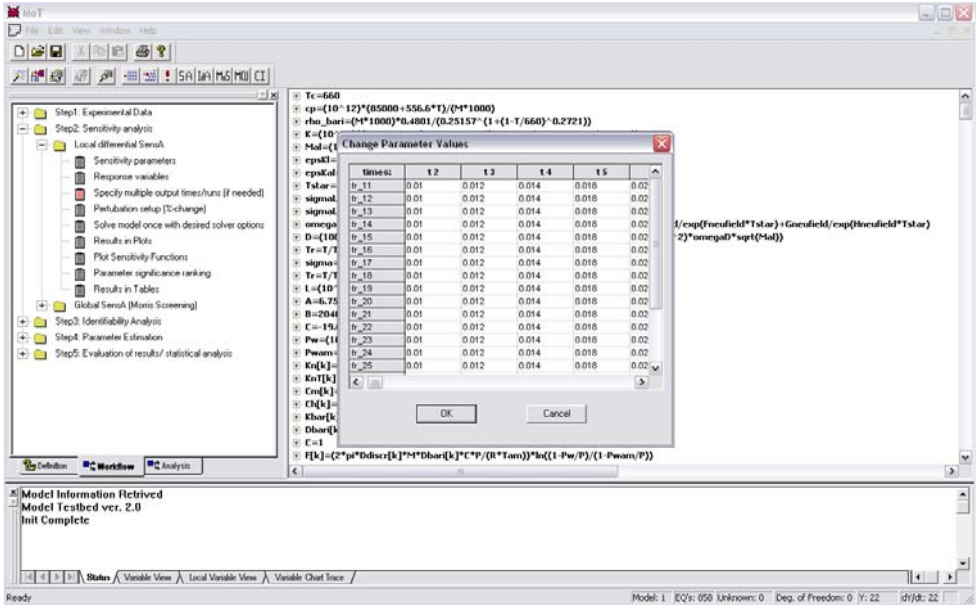


Figure 4.15 *Automated set-up of multiple output times for sensitivity analysis according to experimental data introduced to MoT.*

The local differential sensitivity analysis in MoT also determines a parameter significance ranking for each sensitivity run and output time. The applied sensitivity measure considers the effect on all response variables for each parameter. In addition, an overall parameter significance ranking is determined based on the same sensitivity measure which combines the sensitivity results from all runs, output times and response variables for each parameter. The general equation for the applied sensitivity measure δ_j^{msqr} (Brun et al., 2002; Sin & Vanrolleghem, 2007; Sin et al., 2010) is given by:

$$\delta_j^{msqr} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (S_{nd}(i,j))^2} \tag{4.3}$$

Figures 4.16 and 4.17 show the output of the calculated sensitivities and the parameter significance ranking, respectively.

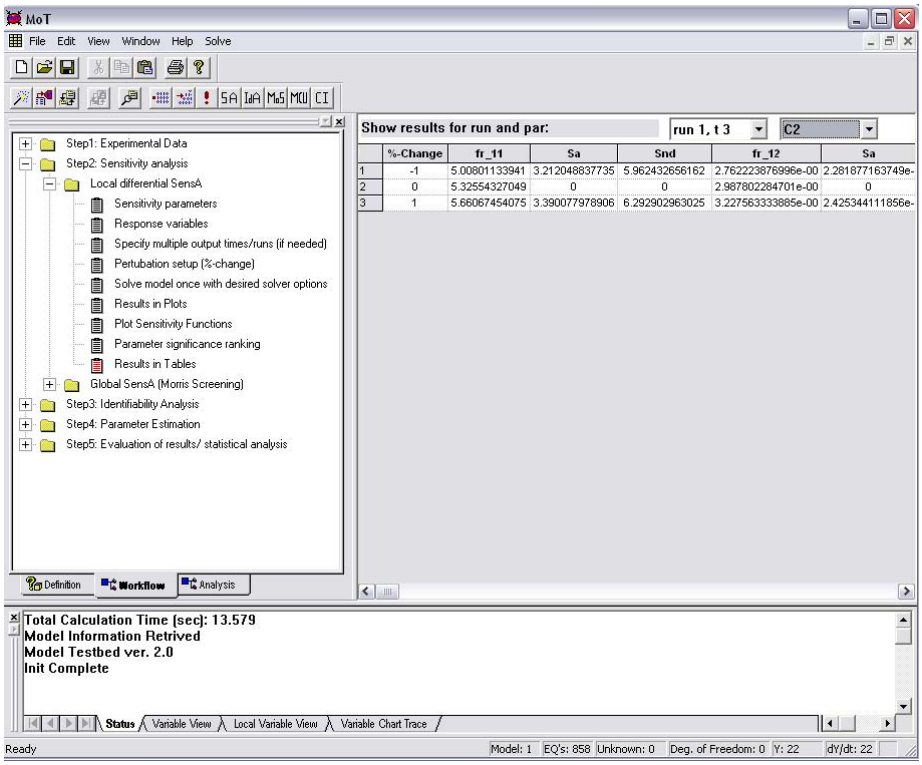


Figure 4.16 *Output of sensitivities in tables.*

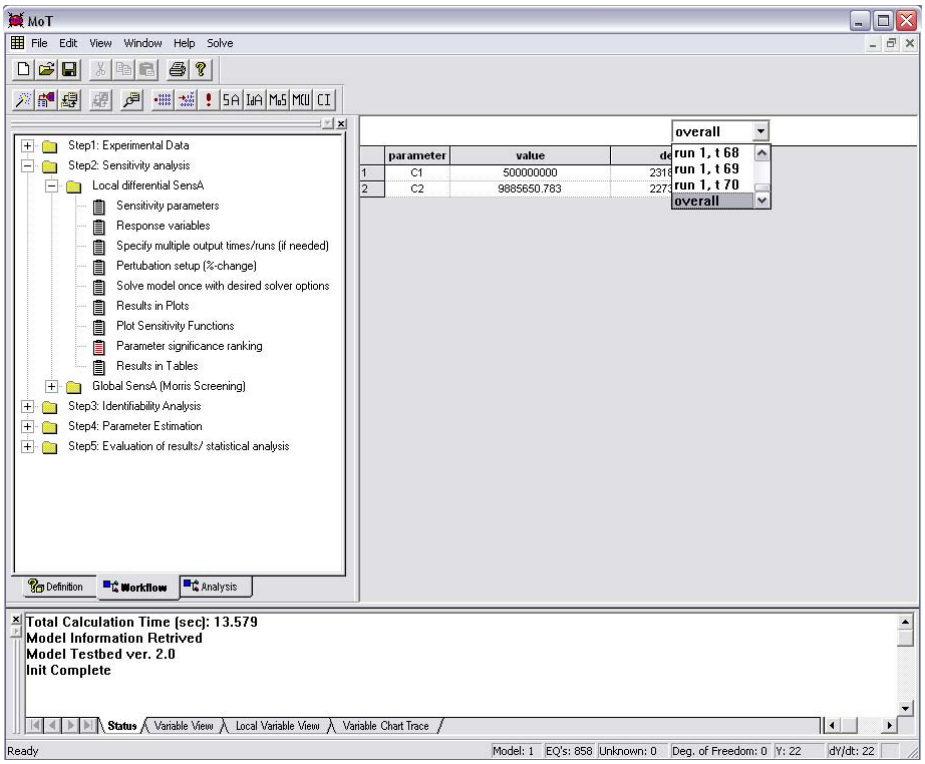


Figure 4.17 *Output of parameter significance ranking in tables.*

The correlation of two parameters is graphically evaluated by plotting the sensitivity functions, which show the normalized sensitivities $S_{norm,ij}$ of the parameters at all datapoints. The normalized sensitivity is calculated as follows:

$$S_{norm} = \{S_{norm,ij}\} \;\; with \;\; S_{norm,ij} = \frac{Snd_{ij}}{\|Snd_{ij}\|} \tag{4.4}$$

If the sensitivity functions are collinear for two parameters for all datapoints these parameters are correlated with respect to the experimental data and cannot be unambigously identified. MoT allows the plot of sensitivity functions for all possible parameter pairs. Figure 4.18 shows the plot of the sensitivity function for two non-collinear parameters.

The global sensitivity method is based on a Morris screening. The Morris screening has been implemented based on Matlab codes provided by Sin et al. (2009b).

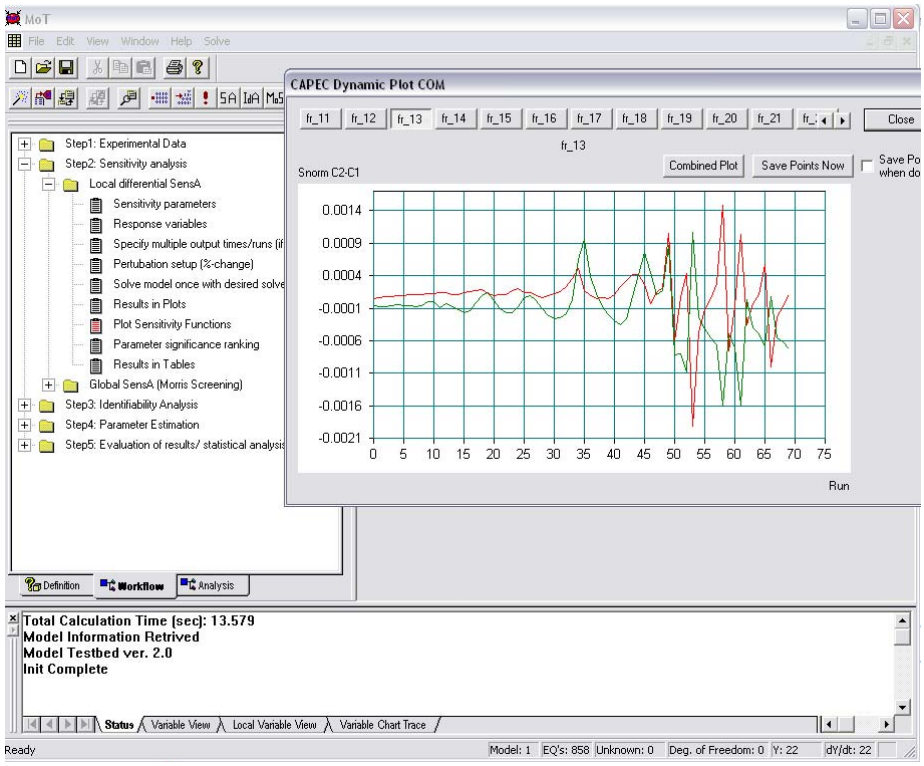


Figure 4.18 *Sensitivity functions for two non-collinear parameters (plotted by MoT).*

Step 3: Identifiability analysis: This step is also optional. Here, the collinearity of all possible parameter subsets is evaluated based on a collinearity index γ_K for the subset K (Brun et al, 2001; 2002; Sin & Vanrolleghem, 2007; Sin et al., 2010):

$$\gamma_K = \frac{1}{\sqrt{\min(\lambda_k)}} \tag{4.5}$$

$$where: \lambda_k = eigen\left(S_{norm,K}^T S_{norm,K}\right) \tag{4.6}$$

Here, $S_{norm,K}$ is the normalized sensitivity matrix of the parameter subset K which is automatically derived by MoT from the sensitivity analysis results for each parameter subset. If the collinearity between the parameters of a subset increases the collinearity index approaches infinity. If the collinearity decreases it approaches unity. In general the collinearity index can be interpreted as follows: The effects of the change of the value of one parameter in a subset K on the response variables can be cancelled out (at least in linear approximation) up to a fraction (given by the ratio $1/\gamma_K$) by adjusting the remaining parameters in the subset (Brun et al., 2001). Consequently, a parameter subset K is considered not identifiable by the available data if its collinearity index exceeds a certain threshold. Thresholds between 10 and 20 have been suggested and applied in literature (Brun et al, 2002; Sin & Vanrolleghem, 2007). Deciding which

threshold value to use depends on the model application purpose (e.g. what level of parameter uncertainties are acceptable) and is usually found after an iterative process (Brun et al., 2001).

MoT also determines a second measure which combines the two criteria for identifiability (sensitivity and non-collinearity), the determinant measure ρ_K (Brun et al., 2002; Sin & Vanrolleghem, 2007; Sin et al., 2010):

$$\rho_K = det\left(S_{nd,K}^T S_{nd,K}\right)^{1/2N_K} = \left(\prod_{j=1}^{K} \lambda_{Snd,j}\right)^{1/(2N_K)} . \quad (4.7)$$

The product of the eigenvalues $\lambda_{Snd,j}$ becomes large if the δ_j^{msqr} values are high and γ_K is low. N_K gives the number of parameters in the subset K. The exponent $1/(2N_K)$ is introduced to the equation in order to allow the comparability of subsets having different numbers of parameters. Since the value of ρ_K depends on the perturbation applied in the sensitivity analysis a general threshold above which a subset of parameters is identifiable cannot be given. Consequently, ρ_K is a relative measure for the comparison of different parameter subsets. Figure 4.19 shows the steps of the identifiability analysis in the MoT interface together with the output table of the subsets that have resulted to be identifiable.

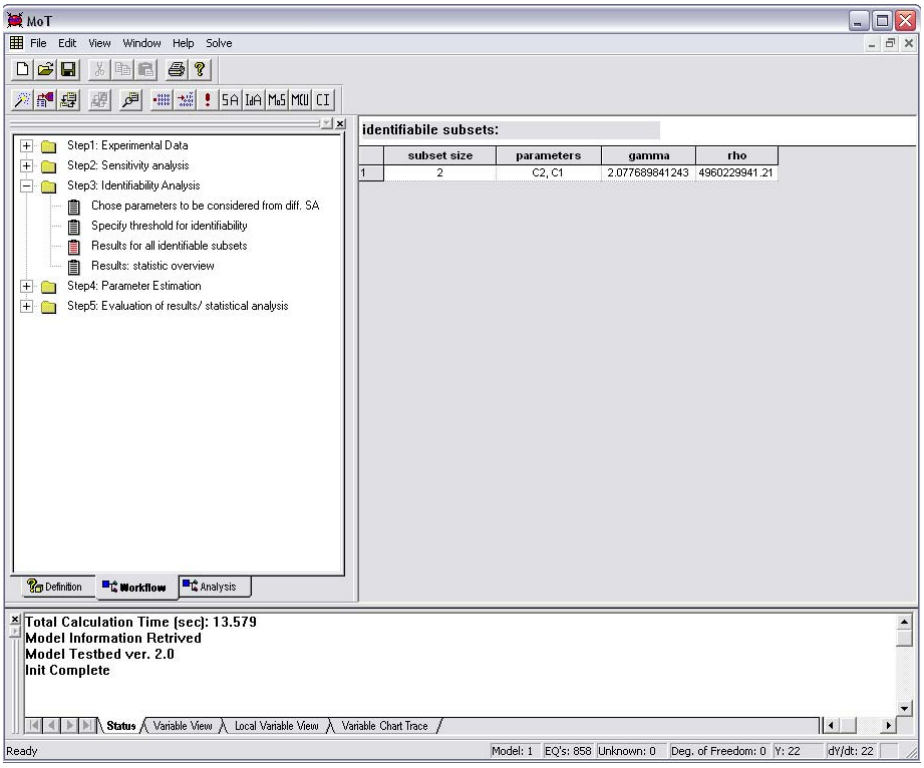


Figure 4.19 *Output of table of identifiable parameter subsets in MoT.*

Step 4: Here, the parameter estimation problem is set up. The set-up is supported by MoT according to the description given in Chapter 3 (Figure 3.4). Figure 4.20 shows a screenshot of the interface for the selection of the objective function including possible weight factors.

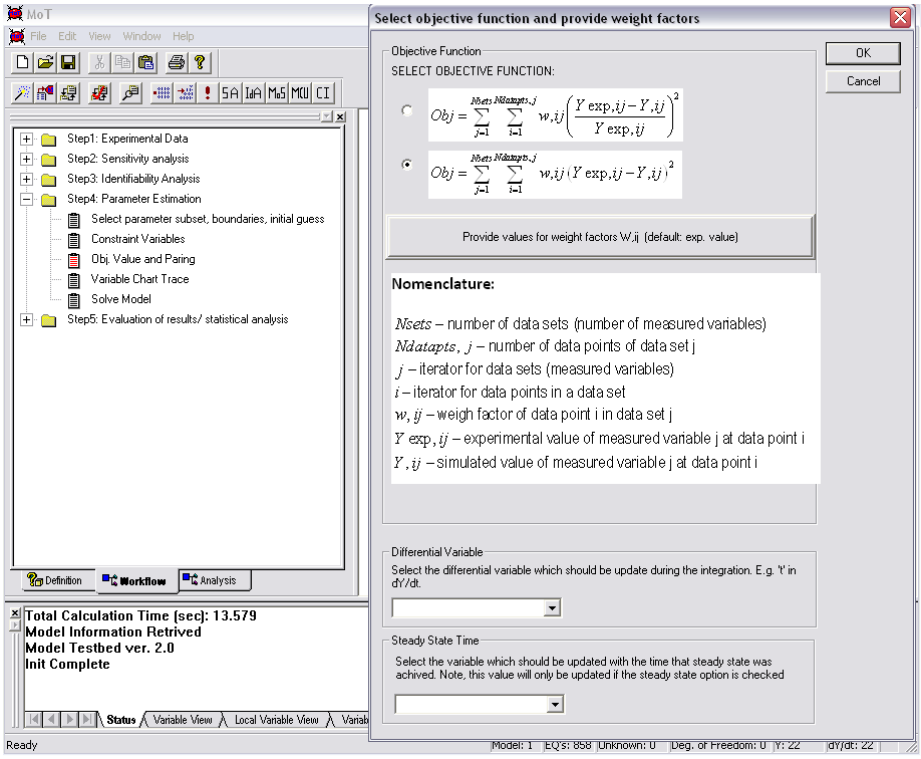


Figure 4.20 *Mot interface for selection of objective function.*

In general it is possible to use least square fit or maximum likelihood estimation (MLE) objective functions in MoT. The MLE objective function cannot be selected and generated automatically but instead needs to be added manually to the model equations. The provided optimizer is based on the SQP algorithm.

Step 5: Here, statistical analysis on the estimated parameter values if performed. According to the requirements identified for the statistical analysis step, MoT provides tools for confidence interval calculation and Monte Carlo uncertainty analysis (with latin hybercube sampling). The confidence interval calculation and the uncertainty analysis have been implemented based on Matlab codes provided by Sin & Vanrolleghem (2007), Sin et al. (2010) and Sin et al. (2009b), respectively. Figure 4.21 shows the different sub-steps of the statistical analysis (left) in MoT and a plot of experimental data compared to simulation results for the estimated parameter values generated by MoT.

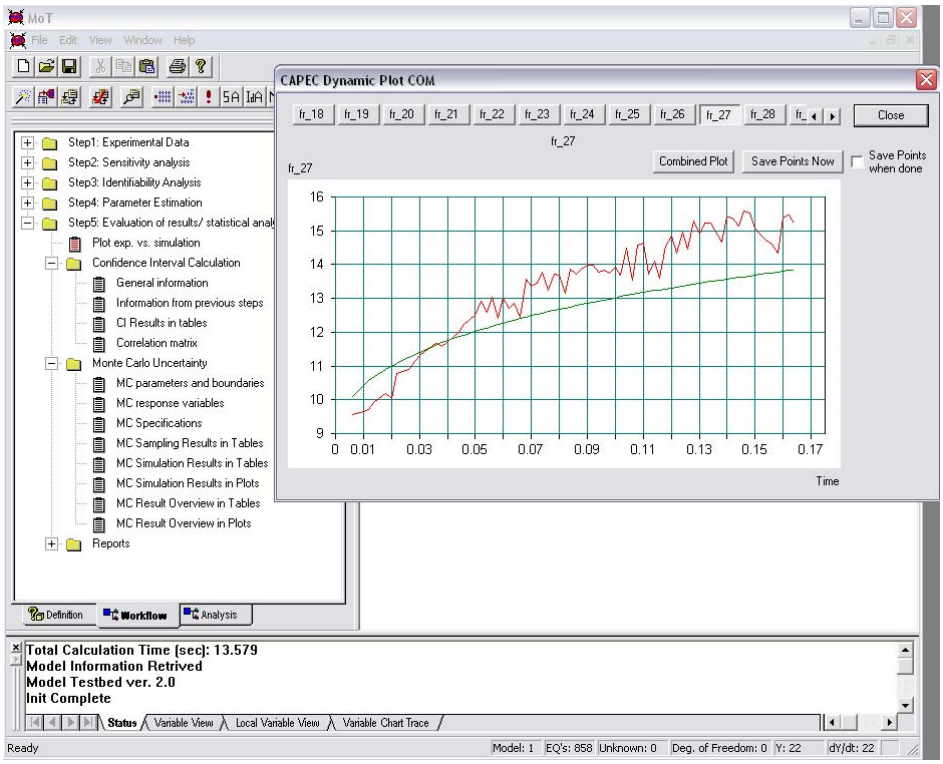


Figure 4.21 *Main interface for evaluation of results/statistical analysis in MoT and generated plot of experimental data vs. simulation results with optimized parameter values.*

4.1.4 Phase IV. Model evaluation/ validation in MoT

Model evaluation only makes sense after the multiple scenario manager has been implemented because otherwise the performance of different model scenarios cannot be compared in the same MoT-file. For this reason, the work-flow for this phase has not yet been implemented in MoT. The evaluation of the model performance is done separately for different candidate models at the end of the model identification work-flow (see previous section, Figure 4.21).

4.1.5 Phase V. Model application in MoT

The two work-flows provided in MoT for model application are the simulation and the optimization work-flows. For the simulation work-flow all features and automation options identified in Chapter 3 (see Figure 3.5) are available, except for support related to the generation of bifurcation curves and the automatic derivation of the corresponding steady state model from a dynamic model.

The optimization work-flow has been implemented exactly as suggested in Chapter 3 including all features and automation given in Figure 3.6. Figure 4.22 shows a screenshot of the

optimization work-flow. Figure 4.23 shows how, after having performed the optimization, the modeller can run a simulation with the optimized design variable values.

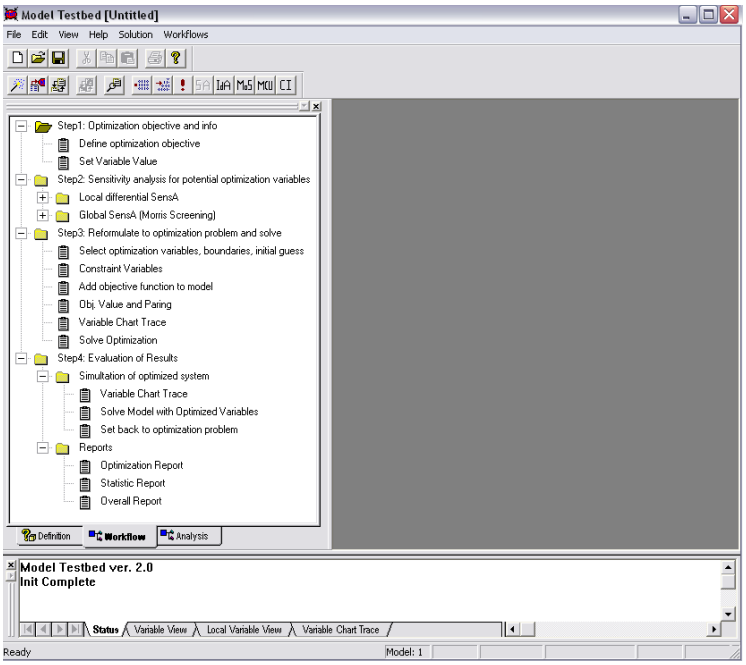


Figure 4.22 *Optimization work-flow in MoT.*

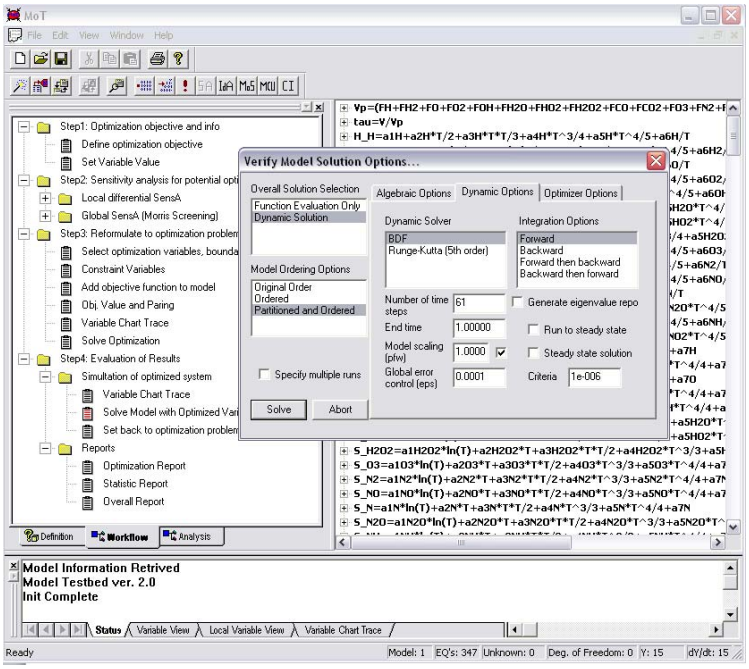


Figure 4.23 *Run simulation with optimized design variable values and generic solver interface in MoT.*

In the background MoT changes the optimization problem into a simulation problem, sets the new design variable values, identifies the solution strategy and required solvers for the problem and opens the generic solver interface where the modeller has the opportunity to select a different solver from the default solver and modify the solver options. In most cases the default settings can be kept. In the dynamic solver options the modeller may wish to adjust the desired final integration time.

4.1.6 Conclusions

From the different MoT-features presented in this Chapter the following features are not available in other existing modelling tools (see Table 1.5).

- Structure based on computer-aided, partly automated work-flows for the different modelling tasks required during model development and application;
- Structured model documentation interface;
- Automated report generation of all performed work-flow steps, structured according to work-flows;
- Partly automated numerical model analysis including display and analysis of incidence matrix.

Although a number of MoT-features are available in some specialized software tools, these tools are lacking the variety of features provided in MoT. For example, the provision of the model equations as close as possible to the way they are written in scientific papers is likewise supported by MOSAIC (Kuntsche et al., 2011). MOSAIC however, does not allow the performance of tasks like model identification, simulation or optimization but instead requires the export of the created model to a different target software. Tools like gProms (Process Systems Enterprise, 2010a-d; 2011) and Aspen Custom Modeler (Aspentech, 2003; 2011a) for example, do not provide features for sensitivity analysis, identifiability analysis and the provision of the model equations is by far more complex than in MoT (e.g. declaration of variables is required , syntax for PDE-discretization more complex).

4.2 Software architecture for extensions of MoT

This section gives a brief overview of some features still missing in MoT and some ideas on how these features could be implemented. Mainly, features for the multi-scale model construction work-flow and a multi-scenario manager have to be developed.

4.2.1 Extension of multi-scale model construction work-flow

The main shortcomings of the current multi-scale model construction work-flow are within the data-flow and linking schemes.

One suggestion is the implementation of templates for specific systems to be modelled and for modelling problems. The idea is that the templates are super-imposed to the work-flows to offer additional specific support and domain knowledge for each step. It needs to be possible to select the templates when the type of the data-flow scheme is selected (Step II.1 of work-flow). More detailed information on how such templates can assist the modeller during model derivation and construction is given in Section 5.4 for the fragrance aerosol case study.

Another potential for improving the computer-aided multi-scale model construction work-flow lies in the improvement of the linking scheme. The linking scheme interface currently does not provide all linking scheme options identified in Chapter 2 (see Step II.4 of the multi-scale model construction work-flow). For variable communication it would be convenient if the modeller could decide if for a communicated vector variable instead of the vector the mean value from all vector elements is to be communicated, or the sum of all elements, or only a boundary element of the vector. In addition it should be possible for the modeller to specify that a variable is to be multiplied by a certain value before communicating it to the other scale.

4.2.2 Implementation of multiple-scenario manager

A multiple-scenario manager allows the storage of different model scenarios for the same modelling problem (single- or multi-scale) in one MoT-file. The MoT-file which contains the different scenarios is called a scenario-file and it allows opening, developing, modifying, comparing and applying the different scenarios. This section communicates some ideas on how this feature could be implemented in MoT.

The scenarios generated through the scenario-file should be also automatically created (and updated) as separate scenarios so that they can be used on a stand-alone basis. The modeller can select the scenario one wants to work on and consequently the corresponding information is loaded to the MoT-interface. It is possible to modify variable values, linking/data-flow schemes, or access the different files for the models of the scenario in order to make changes from the scenario-file. Furthermore, model parameters can be identified or the scenario can be applied for simulation or optimization. It is possible to create reports for the different scenarios and additionally a comparison report that compares the performance of all stored scenarios could also be generated. Figure 4.24 gives a suggestion on how the MoT interface could look like with the implemented multiple-scenario manager feature.

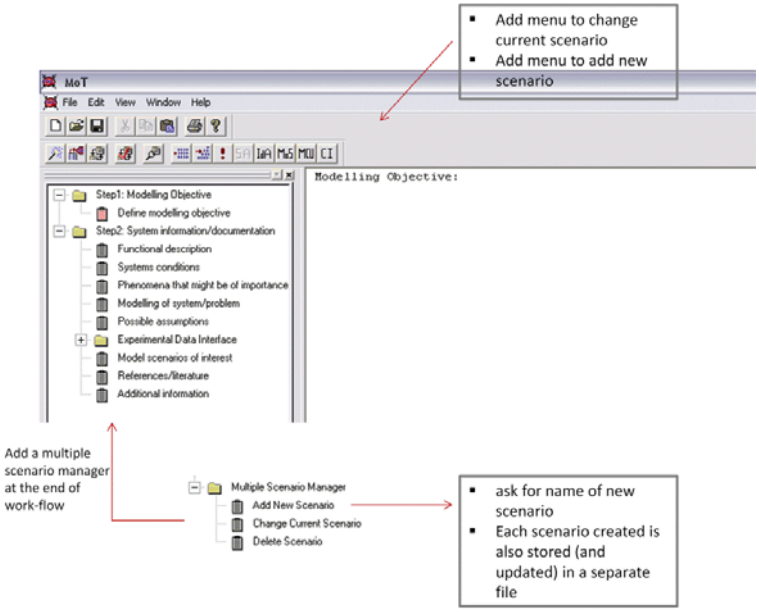


Figure 4.24 *Possible interface for multiple-scenario manager in MoT.*

A scenario manager section can be added at the end of the 'Modelling objective and systems information'-work-flow. From here new scenarios can be added, the current scenario to work on can be changed or scenarios can be deleted. A quick access to add a new scenario or to change the current scenario could be added to the taskbar of MoT (see Figure 4.24).

Chapter 5. Case studies

A number of case studies relevant to different areas and industries in chemical and biochemical engineering have been solved to illustrate and validate the generic work-flows (see Chapter 2), the resulting computer-aided modelling framework (see Chapter 3) and their implementation (see Chapter 4). An overview of the case studies is given in Table 5.1.

Table 5.1 *Overview of case studies*

Case study:	***Related industry***	***Required work-flows***
Thermal treatment of off-gas stream of an adipic acid production process *(Section 5.1)*	Air cleaning industry	Modelling objective and system information; Single-scale model construction; Model identification/ discrimination; Optimization
Batch protein uptake of lysozyme by sepharose beads *(Section 5.2)*	Biotechnology industry	Modelling objective and system information; Multi-scale model construction; Simulation
Fluidized bed reactor *(Section 5.3)*	Chemical industry	Modelling objective and system information; Multi-scale model construction; Simulation
Fragrance aerosol system *(Section 5.4)*	Fragrance industry	Modelling objective and system information; Multi-scale model construction; Model identification/ discrimination; Simulation
Pharmacokinetic modelling of drug uptake and distribution *(Section 5.5)*	Pharma industry	Modelling objective and system information; Multi-scale model construction; Model identification/ discrimination; Simulation

The presented case studies highlight the applicability of the generic work-flows for different modelling problems. Depending on the problem and modelling goal for each case study a different combination of the generic work-flows is needed. The work-flow of Phase I (Modelling goal and system information) is needed in all modelling problems because the modelling objective needs to be formalized as clearly as possible before starting model construction or derivation. All case studies have been solved applying the implemented work-flows in MoT. The features MoT provides for the different work-flow steps have been explained in Chapter 4 and

will not be repeated in detail for the case studies. However, for the last case study, which deals with the development of a pharmacokinetic model for drug distribution in rats (see Section 5.5), the MoT features and the interface are presented together with the work-flow steps. The aerosol fragrance case study demonstrates how a system-specific template can be super-imposed on the generic work-flow steps providing thereby, domain-knowledge and specific support for aerosol fragrance systems.

5.1 Air cleaning industry: Thermal treatment of the off-gas stream of an adipic acid production process

This case study is related to the thermal treatment of the off-gas stream of an adipic acid production process. The objective here is to highlight the general modelling methodology developed in Chapter 2 and its work-flows together with the corresponding methods and tools. The construction and identification of the combustion model is investigated. The developed model is then applied for a reactor design problem. In this way, the case study highlights the 'Modelling objective and systems information'-work-flow (Figure 2.2), the 'Single-scale model construction'-work-flow (Figure 2.3), the 'Model identification/discrimination'-work-flow (Figure 2.6) and the model application work-flow for optimization (Figure 2.8). The case study has been published (Heitzig et al., 2011a) and this chapter is based on this publication.
The model development process is started in Phase I with the 'Modelling objective and systems information'-work-flow:

5.1.1 Phase I. Modelling objective and system information

Step I.1: Modelling objective (Phase I)

The goal for this case study is to identify a model for the thermal treatment of the off-gas stream of an adipic acid production process in a flow reactor/heat exchanger (plug-flow reactor). The model is to be applied for reactor design in order to remove the N_2O (greenhouse gas, source of O_3 in stratosphere). Consequently, the model needs to be able to calculate the N_2O outlet concentrations of the reactor at different temperatures and residence times/reactor volumes.

Step I.2. System information (Phase I)

Figure 5.1 shows a basic sketch of the thermal treatment of the off-gas stream of the adipic acid production.

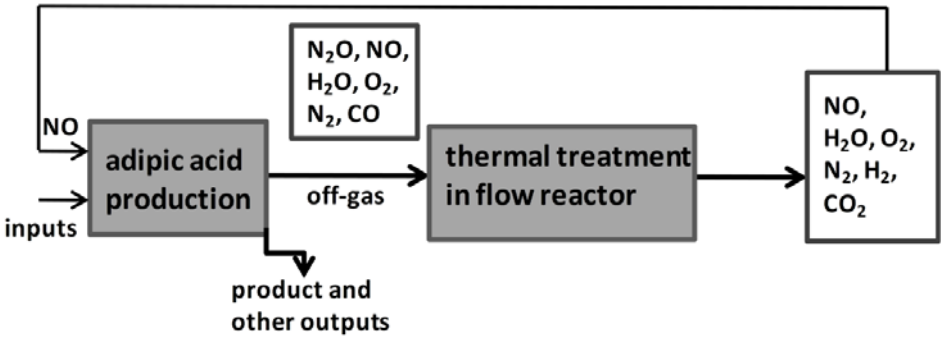


Figure 5.1 *Thermal treatment of adipic acid production off-gas stream.*

The system under consideration (thermal treatment unit) consists of a total of 18 compounds. Many of these compounds occur in very low concentrations. Figure 5.1 shows the important input and output compounds. NO is recycled back to the adipic acid production unit because it is a feedstock material for one of the production steps. For the H_2/O_2 combustion mechanism information involving 18 elementary, reversible reactions (rate constants, thermodynamic properties) is available from Rasmussen et al. (2007). For the nitrogen species and their reactions, information is initially taken from the NIST Chemical Kinetics Database (National Institute of Standards and Technology, 2000). In total there are 44 reactions in the system. The rate constants for the forward reactions are calculated applying the Arrhenius equations. The backward rate constants can be calculated from the forward rate constants and the equilibrium constants $K[k]$. In order to calculate the equilibrium constants the compound standard enthalpies $H_j^O(T)$ and entropies $S_j^0(T)$ are needed. The Nasa polynomials (CHEMKIN Collection Release 3.6, 2000; Kee et al., 1994) provide correlations for $H_j^O(T)$ and $S_j^0(T)$ with respect to the reactor temperature for all components in the system. For pressure dependent reactions, such as the dissociation of N_2O, the third body enhancement needs to be considered. Third bodies are molecules which promote the reaction but remain chemically inert during the reaction. The behaviour of the rate in the fall-off regime can be calculated applying the Troe equation (Troe, 1979).

The expected operation conditions for the thermal treatment are a pressure of 1 atm due to safety and economic reasons and an arbitrarily chosen maximum temperature of 1500 K, which is due to material limitations.

Experimental data to identify the model parameters are taken from Glarborg et al. (1994) and given in Appendix A2. The data involve 76 data-points and are subdivided into five datasets. For each dataset the feed concentrations as well as the residence time differ, whereas the pressure is constant at 1.05 atm for all datasets. The measured variable is the output concentration of N_2O for different temperatures.

5.1.2 Phase II. Model construction

The reactor model is constructed based on the information collected in the previous step applying the single-scale model construction work-flow (Phase II.A).

Step II.1: Model scenario documentation and concept (Phase II)

A number of assumptions are made for the chemical system. It is assumed to be ideally mixed in the radial reactor direction. The reactor is modelled as a plug flow reactor. The system is assumed to be isothermal and isobaric. Transport phenomena like diffusion and dispersion are neglected. Furthermore, the system is considered to be at steady state. The assumptions of an ideal reactor may not be fulfilled in a practical system, but they are often used in chemical engineering as they serve to simplify the model analysis (Zwietering, 1959). Evaluation of possible temperature and velocity gradients is outside the scope of the present work.

The phenomena considered in the model are the convective mass transport along the reactor axis and the kinetics of the chemical reactions in the system. The model description corresponding to the made assumptions and considered phenomena is summarized in Figure 5.2.

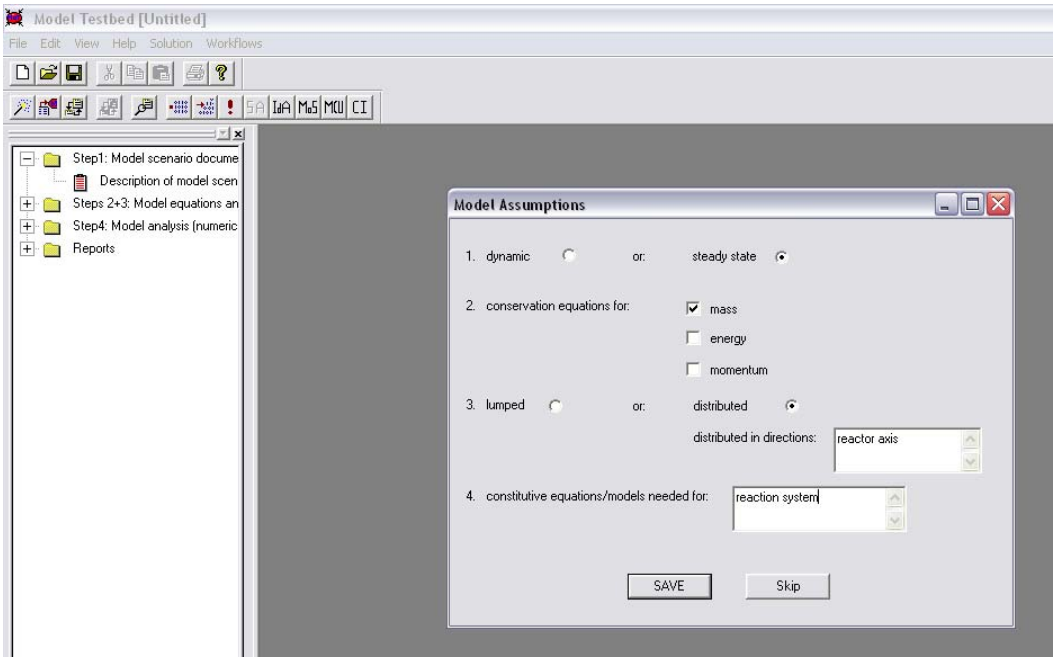


Figure 5.2 *Model description.*

Steps II.2+3: Derivation of model equations and model translation (Phase II)

The complete set of model equations is given in Appendix A3. Due to the assumptions that the system is isothermal and isobaric no energy and momentum balances are needed. Mass balances are required for the 15 non-inert compounds in the system. Since the system is considered to be distributed in the axial reactor direction the model equations need to be discretized accordingly. Here, they are transformed to a system of ordinary differential equations (ODEs) (1 ODE for each non-inert compound) having the reactor length as independent variable. This is possible due to the steady state assumption for the system. Constitutive equations are required to provide expressions for the reaction rates appearing in the mass balance equations. The derived model equations have been introduced to MoT and then translated.

Step II.4: Model analysis (numerical) (Phase II)

The first step after the construction of the model is to determine the number and types of equations. The system has 15 ordinary differential equations (ODEs) and 330 algebraic equations (AEs). The 613 model variables can be pre-classified into 15 dependent variables (the flow rates of the non-inert compounds), 1 independent variable (the reactor volume) and 598 algebraic variables. The degrees of freedom for the algebraic equation part are 268. They are obtained as the difference between the number of algebraic variables appearing in the AEs and the number of AEs. Accordingly, 268 variables need to be specified either as parameters or known variables. Table 5.2 gives an overview of the specified variables.

Table 5.2 *Specified variables*

variable type	***variables***	***number***
parameter	coefficients for the Arrhenius equations	144
	coefficients of the Nasa polynomials	105
	Troe equation parameter	13
		total 262
known	molar flows of inert compounds	3
	pressure	1
	temperature	1
	gas constant	1
		total 6

During this specification step a singularity check has been conducted. The remaining 330 algebraic variables are unknown (Table 5.3). For this case study all unknown variables are explicit variables.

Table 5.3 *Unknown variables*

variables	***number***
volumetric flow	1
component enthalpies	15
component entropies	15
reaction enthalpies	44
reaction entropies	44
equilibrium constants	44
3rd body concentrations	7
rate constants for forward reaction	44
rate constants for backward reactions	44
Troe equation variables	28
reaction rates	44

The degrees of freedom of the ordinary differential equation part equal the number of algebraic variables in the ODEs which do not appear in the AE part. In this case it is 0.

The next step is to generate the incidence matrix and based on that the equations are ordered (Table 5.4). The equations are grouped into algebraic and ordinary differential (last row of incidence matrix) equations. Since the incidence matrix of the system is rather big (345x345) it is shown in a 'condensed' form where the equations and variables are represented by vectors, e.g. all component enthalpies are represented by the vector $\mathbf{H}_j^0$. The incidence matrix reveals that the AEs and ODEs are coupled (off-diagonal elements). Apart from the coupling to the ODE part the algebraic equations are all explicit (lower tridiagonal form). For this reason, it is easy to solve the AE and ODE parts separately since at any time t, the value of $\mathbf{F}_j$ would be known, and the AEs can be solved in the sequence given in Table 5.4.

Table 5.4 *Incidence matrix after equation ordering*

	Vp	H_j^0	S_j^0	HR_k	SR_k	K_k	F_{Mk}	k_{infk}	k_{lowk}	X_k	F_{centk}	c_k	N_k	F_k	kf_k	kb_k	r_k	F_j
*(A3.1)**	*																	⊛
(A3.2-16)		*																
(A3.17-A3.31)			*															
(A2.32-A3.75)		*		*														
(A3.76-A3.119)			*		*													
(A3.120- 163)				*	*	*												
(A3.204-210)							*											⊛
(A3.211-214)								*										
(A3.215-218)									*									
(A3.219-222)	*						*	*	*	*								
(A3.223-226)											*							
(A3.227-230)											*	*						
(A3.231-234)											*		*					
(A3.235-238)										*	*	*	*	*				
(A3.164- 203, 239-242)															*			
(A3.243-286)						*									*	*		
(A3.287-330)	*														*	*	*	⊛
(A3.331-345)																	*	*

* *Equation number according to Appendix A3*

After having performed the previous analysis steps the variables classified as known variables need to be given a value. Also for the parameters a value or an initial guess (if they are to be identified by experimental data in a later step) needs to be provided. Further, initial conditions for the dependent variables are required- in this case for the compound flows at the entrance of the reactor.

Before proceeding to the next step, the eigenvalues of the system have been determined for the conditions of dataset 2 at a temperature of 1381 K during a simulation for 646 time-steps. The eigenvalue analysis indicates whether the system would converge to an asymptotically stable steady state (for the investigated conditions) or not. Furthermore, they give information on possible stiffness, oscillations and potential for model reduction. The stiffness ratio is defined as the quotient of the maximum absolute value of the real parts and the minimum absolute value of the real parts of the eigenvalues. For the first time-step it results to be 8.35×10^{20} whereas the stiffness ratio for the last time-step is 2.09×10^{19}. From this it can be concluded that the system is highly stiff having very fast modes on the one hand and slow modes on the other. Since the real parts of all eigenvalues are negative the system can be stated locally (asymptotically) stable.

Furthermore, there are no oscillations. A dynamic solver that can handle the stiffness of the system needs to be selected. In this case the BDF-solver provided in MoT has been used.
The construction of the model is now completed. In a next step the model identification work-flow is used in order to fit the model parameters using the available experimental data (Appendix A2).

5.1.3 Phase III. Model identification

The model identification work-flow has been applied.

Step III.1: Experimental data (Phase III)

The experimental data given in Appendix A2 is used for model identification. It consists of five different measurement sets, which differ in their initial feed composition and residence time. Each set contains 13-15 measurements of the steady state concentration of N_2O at the reactor outlet for different reactor temperatures. N_2O is also the concentration of interest with respect to the modelling goal (see Phase I). The reactor was operated under isothermal conditions. The total number of data-points is 76.

Step III.2: Sensitivity analysis (Phase III)

A local differential sensitivity analysis is conducted in order to ensure that measured variables are actually sensitive to the parameters to be estimated (for more details on the method see Section 4.1.3). The analysis needs to be conducted at each data-point available, for all measured variables and with respect to all parameters to be identified. Consequently, NDAT=76 sensitivity runs are needed, each run differing in the feed conditions, residence time and reactor temperature. The thermodynamic properties are considered to be known and therefore only the NPAR=157 kinetic parameters of the model need to be estimated and hence considered in the sensitivity analysis. Consequently, a total of 11 932 (=NDATxNPAR) local differential sensitivity analysis steps (forward and backward) are required for this case study.
It is important to choose a reasonable value for the perturbation. On the one hand, the perturbation of a parameter needs to be small enough so that the forward and backward perturbations cause the same change to the model output. On the other hand, it has to be taken into account that the solver is still able to resolve the effect of a small perturbation. Care was taken to ensure that these criteria were fulfilled for all parameters and data points investigated. The resulting value for the perturbation is 0.01%. Table 5.5 shows the overall parameter significance ranking for the top 20 most sensitive parameters (taking into account all sensitivity runs) based on the sensitivity measure δ_j^{msqr} given in Equation 4.3 (Section 4.1.3).

Table 5.5 *Parameter significance ranking based on sensitivity measure* δ_j^{msqr} *considering the available measurements (perturbation +/- 0.01% of the initial parameter value)*

rank	*parameter*	*parameter value*	δ_j^{msqr}
1	E(37)[1]	62796.0	27.063
2	β(37)[1]	-0.73	6.288
3	A(37)[1]	7.23E+017	1.190
4	β (40)	-2.87	0.825
5	E(43)	15937.0	0.757
6	E(39)	15103.0	0.581
7	A(43)	3.69 E+012	0.131
8	A(39)	9.64 E+013	0.105
9	β (12)	-2.00	0.103
10	E(38)	26629.0	9.39E-002
11	E(1)	16600.0	7.62E-002
12	A(40)	4.71E+024	3.97E-002
13	β (1)	-0.41	3.73E-002
14	E(40)	1552.0	2.23E-002
15	β (32)	-2.16	2.03E-002
16	E(32)	37161.0	1.75E-002
17	A(1)	3.550E+015	1.26E-002
18	β (13)	1.52	9.8E-003
19	A(38)	6.62 E+013	9.6E-003
20	β (36)	4.72	8.9E-003

[1]*A(37), β (37) and E(37) represent the kinetic parameter of the Arrhenius equation for the reaction number 37*

Since only sensitive parameters with respect to the measured variables can be identified, the non-sensitive parameters are fixed to their initial values from literature. The minimum sensitivity measure for a parameter to be still considered in the analysis has been selected to be 10^{-1} based on the previous experiences with the method (Sin & Vanrolleghem, 2007). As a result, only the top nine parameters in the ranking are deemed significantly sensitive and hence considered for further identifiability analysis.

It has been highlighted in Section 4.1.3 that, in addition to the parameter significance ranking, the output of the sensitivity functions during the sensitivity analysis step can be used to further evaluate the identifiability of the unknown model parameters. The sensitivity functions allow to graphically evaluate the correlation between two parameters. The normalized sensitivities $S_{norm,ij}$ (see Section 4.1.3, Equation 4.4) of the parameters at all datapoints (from all data sets) are plotted. If the sensitivity functions are collinear for two parameters they are correlated and cannot be unambigously identified by the available experimental data. The sensitivity functions for all possible parameter pairs from the top 9 most sensitive parameters (still considered after

the sensitivity analysis) are given in Appendix A4. As an example, Figure 5.3 shows the plots of the sensitivity functions for dataset 5 and the 2-parameter subset with the highest (right) and the lowest (left) collinearity index (calculated later, during the identifiability analysis in Step 3). It can be seen that the curves for the subset with the highest collinearity index are collinear whereas this is not the case for the second subset.

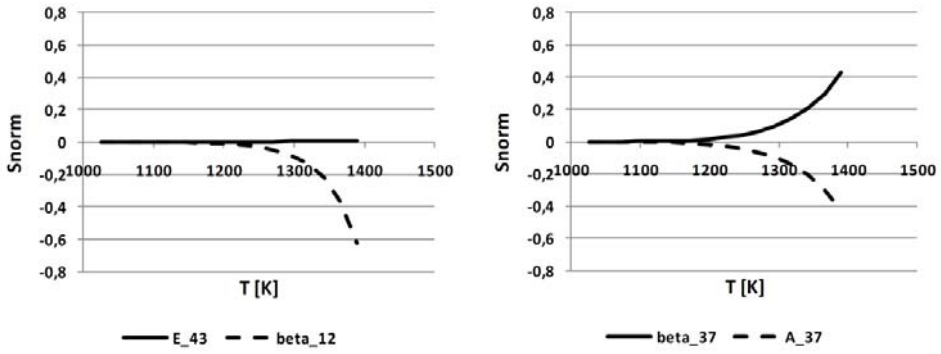


Figure 5.3 *Plot of sensitivity functions for parameter pair with lowest (left) and highest (right) collinearity index (collinearity index determined in Step 3).*

Step III.3: Identifiability analysis (Phase III)

The second condition for identifiability of a set of parameters, next to the required sensitivity, is that there is no collinearity between the parameters. The sensitivity functions obtained in the previous sensitivity analysis step already allowed to graphically evaluate the correlation between the different parameter pairs. It is however recommendable to perform an identifiability analysis based on a quantitative measure which can also consider collinearity between more than two parameters. For this application example it should be mentioned that the Arrhenius parameter for one reaction usually are correlated.

The identifiability analysis has been performed (see Section 4.1.3). In short, all possible parameter subsets from the 9 selected parameters (after sensitivity analysis) need to be generated and each parameter subset K is evaluated based on its collinearity index γ_K (given by equations 4.5 and 4.6). For this case study, a threshold of 10 has been applied for the collinearity index based on the previous experiences with the method. A parameter subset K with a collinearity index γ_K below this threshold is considered identifiable. Table 5.6 gives an overview of the results. From the 502 possible parameters subsets (from the top 9 most sensitive parameters considered in this step) only 134 result to be identifiable, that means, having a γ_K lower than 10. These non-identifiability issues are commonly encountered in engineering models (Brun et al., 2002; Sin et al., 2009a) a discussion of which is beyond the

scope this work. Table 5.6 moreover provides the determinant measure ρ_K which is, next to the sensitivity measure δ_j^{msqr} and the collinearity index γ_K, the third measure for identifiability. The determinant measure ρ_K combines the two previous measures. It is calculated based on Equation 4.7 (in Section 4.1.3).

Table 5.6 *Summary of the identifiable analysis results*

subset size:	***total no. of subsets:***	***not identify-able subsets:***	***% of identify-able subsets:***	***minimal γ_K:***	***maximal γ_K:***	***ρ_K of min. subset:***	***ρ_K of max. subset:***
2	36	31	86.1	1.0	66.3	0.0040	62.26
3	84	51	60.7	2.0	21534.9	0.0054	11.64
4	126	40	31.77	3.8	21739.0	0.0043	3.33
5	126	12	9.5	5.1	25482.0	0.0051	0.87
6	84	0	0	79.1	25770.2	0.0051	0.53
7	36	0	0	145.4	27847.2	0.0069	0.11
8	9	0	0	29947.1	29947.1	0.0063	0.006
9	1	0	0	31127.0	31127.0	0.0048	0.005

Table 5.6 shows that identifiable subsets can only be found up to a subset size of 5 parameters. From a subset size of 2 to a size of 5 parameters the percentage of identifiable subsets decreases from 86% to 9.5%. The maximum collinearity indexes increase with the size of the subsets which means that the collinearity increases with the subset size. The minimal and maximal values for the determinant measure decrease with increasing subset size. This is to be expected since with increasing subset size the collinearity increases and the number of less sensitive parameters that have to be included into the subset increases. Based on the results of the identifiability analysis proper subsets of parameters can be chosen to be passed to the parameter estimation step.

Step III.4: Parameter estimation (Phase III)

The least square fit (objective function) is used for the parameter estimation:

$$Obj = \frac{1}{NDAT}\sum_{i=1}^{NDAT}\left(F_{N_2O}(i) - F_{N_2O}^{exp}(i)\right)^2 \quad (5.1)$$

Here, $F_{N_2O}(i)$ is the model prediction for the volume flow of N_2O at the reactor exit at the conditions of data point i while $F_{N_2O}^{exp}(i)$ is its corresponding measured value. The value of the objective function before regression, when all parameters are set to their initial values from literature, is Obj=905.67. The parameter regression is performed with respect to all 76 available experimental data points at once applying the SQP solver of MoT. The parameter boundaries are set to +/-20% from the initial parameter values from literature. If the regression is performed for all top 9 most sensitive parameters at once (ignoring the identifiability analysis results) the

obtained value of the objective function is 144.67. The resulting parameter values are given in Table 5.7.

Table 5.7 *Results of parameter regression for top 9 most sensitive parameters*

parameter:	***initial guess:***	***final estimated parameter value:***
β (12)	-2.00	-2.33
A(37)	7.23e+017	6.64e+017
β (37)	-0.73	-0.65
E(37)	62796	62322.85
A(39)	9.64e+013	11.57e+013
E(39)	15103.00	12082.40
β (40)	-2.87	-2.30
A(43)	3.69e+012	3.81e+012
E(43)	15937.00	13228.76

In a second step, one of the largest identifiable subsets (5 parameters) is chosen for estimation. Among the twelve identifiable 5-parameter subsets, the subset which has the lowest collinearity index $\gamma_K = 5.11$ is selected. If the initial parameter values are set to the values from the first regression step and only the 5 non-correlated parameters are re-fitted the objective function value can be improved to 120.49. Table 5.8 shows the improved values for the 5 parameters in the subset.

Table 5.8 *Re-identification of identifiable 5-parameter subset with highest* $\boldsymbol{\gamma_K}$

Parameter:	***before estimation:***	***after estimation:***
β (12)	-2.33	-2.61
A(37)	6.64e+017	not estimated
β (37)	-0.65	not estimated
E(37)	62322.85	62409.45
A(39)	11.57e+013	not estimated
E(39)	12082.40	9665.92
β (40)	-2.30	-1.84
A(43)	3.81e+012	not estimated
E(43)	13228.76	12390.03

Step 4: Statistical analysis of model predictions and evaluation (Phase III)

This step evaluates how good the parameters fit the experimental data. Two measures for the quality of the fit have been calculated ('Mean Absolute Error' and 'Root Mean Square Deviation'):

$$MAE = \frac{1}{NDAT}\sum_{i=1}^{NDAT} abs\left(F_{N_2O}(i) - F_{N_2O}^{exp}(i)\right) = 8.311 \ (20.91) \tag{5.2}$$

$$RMSE = sqrt\left(\frac{1}{N}\sum_{i=1}^{NDAT}\left(F_{N_2O}(i) - F_{N_2O}^{exp}(i)\right)^2\right) = 10.977 \ (30.09) \tag{5.3}$$

The numbers in brackets are the corresponding measures resulting from the initial parameter values from literature. The quality of the fit has been improved significantly. Figure 5.4 shows the simulation results in comparison to the experimental measurements applied for model identification.

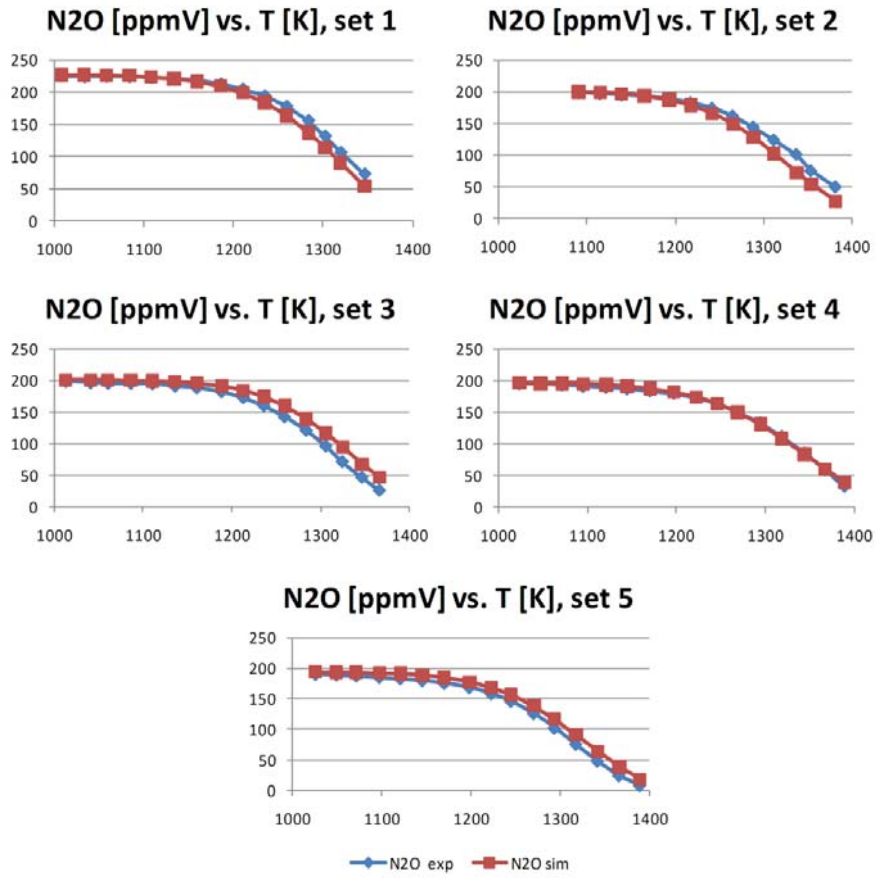


Figure 5.4 *Plot of experimental measurements (Glarborg et al., 1994) and simulations of N2O [ppmV] concentration at reactor exit vs. temperature for the 5 different data sets applied for parameter estimation.*

The figure reveals that the data is represented quite well. Overall the model performance has been decided to be satisfactory.
In a next step the developed model is applied for a reactor design problem. To this end the application work-flow for optimization is applied.

5.1.4 Model application B. Optimization

Now the constructed and identified model is to be used for a reactor design problem, which was the objective at the outset of the modelling study. This means that the optimal design is found by fixing the parameters identified in the previous steps and varying the design variables. In case the design target cannot be met the process concept needs to be revised. The solution of the reactor design problem is performed according to the optimization work-flow (Chapter 2, Figure 2.8) and is described briefly in the following steps.

Step B.1: Optimization objective

For the investigated problem the design target is to reduce the concentration of N_2O at the reactor exit below 100 ppm. At the same time the NO concentration should be maximized since it is an intermediate product for the adipic acid production and can be recycled (see Figure 5.1).
Potential design variables are the temperature, the pressure and the residence time. It has been decided to run the reactor model to steady state during the optimization and therefore the reactor volume is eliminated as a potential design variable.

Step B.2: Update variable values

The feed conditions for the thermal treatment unit differ from that of the available experimental data: 30% N_2O, 0.7% NO, 300 ppm CO, 3% H_2O, 4% O_2, balance N_2.

Step B.3: Derive objective function

The applied objective function (Equation 5.4) combines the two design objectives: the main objective to reduce the N_2O concentration F_{N_2O} below a given boundary and the second objective to increase the NO concentration F_{NO} at the reactor exit.

$$Obj = min\left\{10^6 \cdot F_{N_2O} + \frac{0.1}{F_{NO}}\right\} \tag{5.4}$$

The weighting factors in the objective function have been chosen such, that the impact of both terms in the objective function is of the same order of magnitude for the initial estimates of the design variables.

Step B.4: Sensitivity analysis for design variables

The main motivation for the sensitivity analysis performed here is to get an impression of how strong the impact of the two design variables (pressure and temperature) is on the objective function and if these variables have an opposite impact on the two design objectives combined in the objective function. The base value of the design variables T and P are set to 1450 K and1 atm, respectively. A local differential sensitivity analysis is conducted perturbing each design variable in 5%-steps between a range of -15% and +15%. The response variables are the objective function of the optimization problem (Equation 5.4) as well as the N_20 and NO concentrations. Figures 5.5 and 5.6 show the results for the pressure and the temperature, respectively.

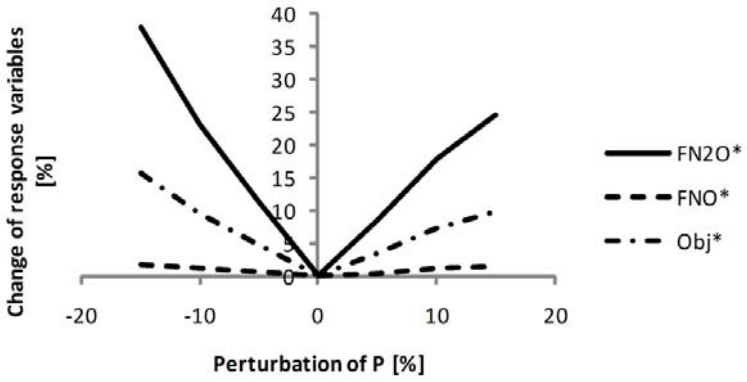


Figure 5.5 *Change of response variables [%, absolute value] versus perturbation of pressure P [%].*

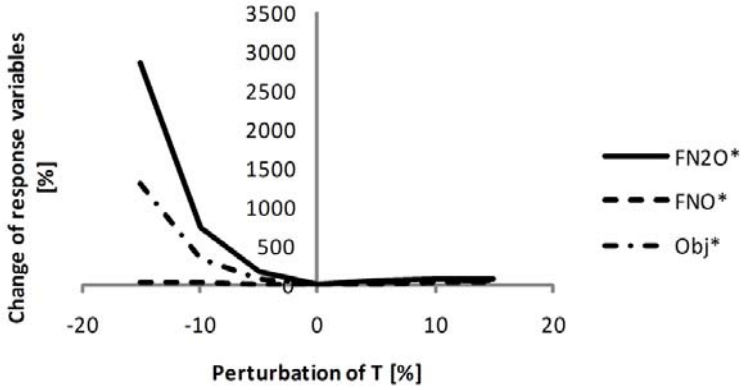


Figure 5.6 *Change of response variables [%, absolute value] versus perturbation of temperature P [%].*

The figures reveal that the molar flow of N_2O is very sensitive to both design variables. However, the sensitivity is remarkably higher for the design variable T. Therefore, optimizing the temperature has a better potential of improving the value of the objective function and the main attention should be paid to this variable during the design process.

Figure 5.7 shows the values of the response variables N_2O, NO and Obj with respect to the temperature T. Figure 5.8 shows the values of the response variables with respect to the design variable pressure P.

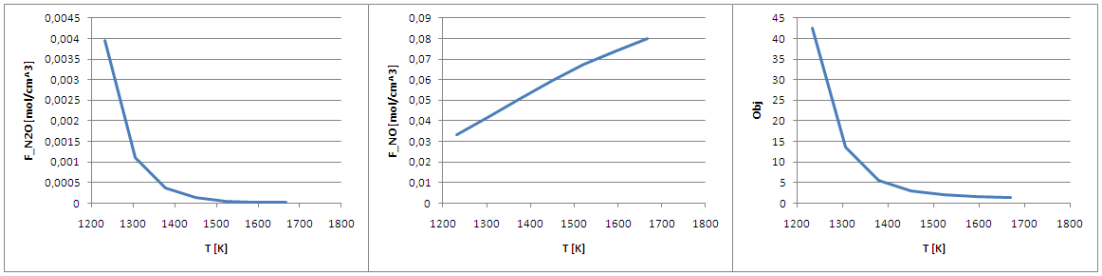


Figure 5.7 *Response variables of sensitivity analysis versus value of perturbed design variable temperature T [K].*

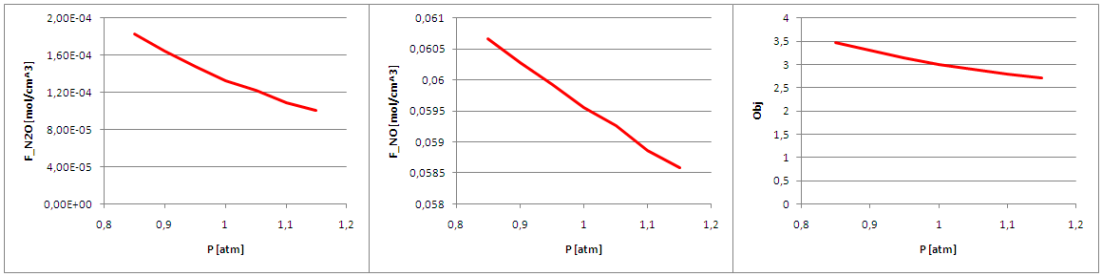


Figure 5.8 *Response variables of sensitivity analysis versus value of perturbed design variable pressure P [atm].*

For the design variable temperature it can be stated that an increase of temperature is favourable for both design objectives, since the NO concentration is increased and the N_2O concentration is decreased at the same time. For the perturbation variable pressure the situation is different. On the one hand, the concentration of N_2O decreases with increasing pressure which is in agreement with the optimization target. On the other hand, the concentration of NO likewise decreases with increasing pressure which is not desired.

Step B.5: Re-formulate to an optimization problem and solve

The problem needs to be re-formulated to an optimization problem by adding an appropriate objective function and transforming the model equations to constraints:

$$Obj = min\left\{10^6 \cdot F_{N_2O} + \frac{0.1}{F_{NO}}\right\} \tag{5.5}$$

$$s.t.: model\ equations$$

$$T \leq 1500\ K$$

The temperature constraint in Equation 5.5 is due to material limitations. Apart from that, due to safety and economic issues, pressures different from the atmospheric pressure are only acceptable if they lead to significant improvement of the reactor performance. The boundaries for the design variables are set to [280 K, 1500 K] for T and to [0 atm, 3 atm] for P. The initial values given to the design variables are T=1450 K and P=1 atm. The applied optimization method is SQP with a convergence criterion for the normalized step length of 10^{-10}. Figure 5.9 shows the values of P and T during the solution of the optimization problem given in Equation 5.5.

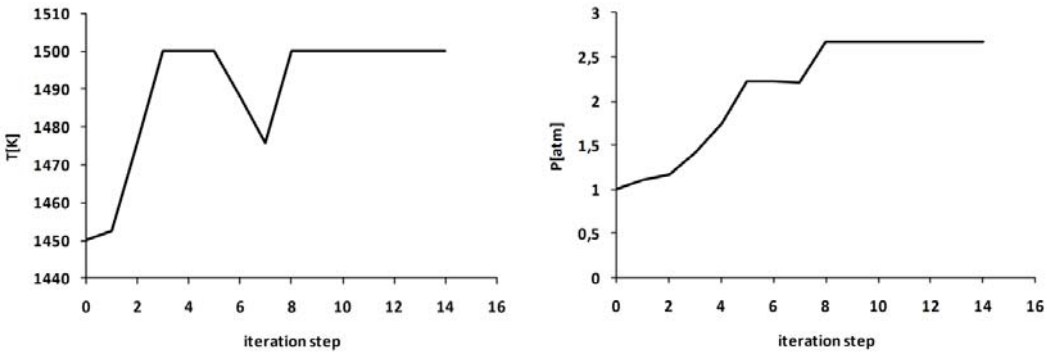


Figure 5.9 *Optimization of design variables T[K] and P[atm] versus iteration steps.*

Figure 5.10 provides a surface plot of the objective function around the found optimum. The best value of the objective function is obtained for a temperature of 1500 K and a pressure of 2.67 atm. It turns out, however, that if a pressure of 1 atm is applied, the N_2O concentration is higher but does not exceed the maximum allowed value. Consequently, a pressure of 1 atm is chosen which is economically favourable, has advantages for the construction of the reactor and at the same time increases the NO concentration at the reactor exit.

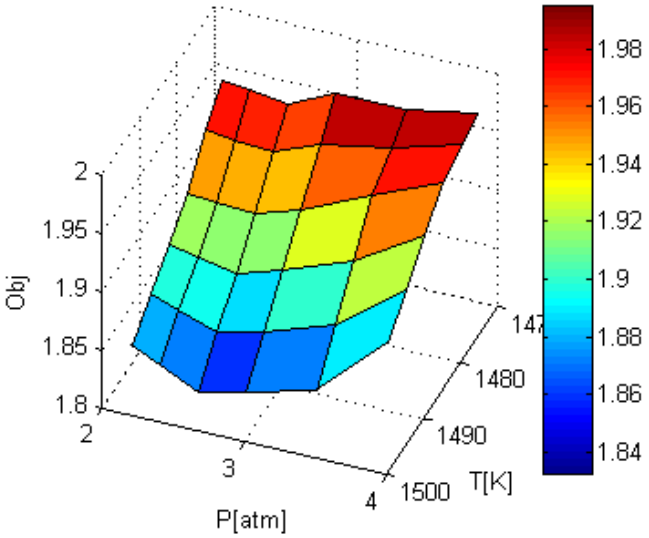


Figure 5.10 *Surface plot of objective function during optimization of design variables P [atm] and T [K].*

Step B.6: Simulation of system with optimized design variable values

In this last step, simulations are performed for the optimized design parameters in order to study the performance of the system under these conditions. The final concentrations of N_2O and NO for a temperature of 1500 K and a pressure of 1 atm are 1.24 ppm and 0.065 [mol/mol], respectively. The boundary value of for the maximum allowed N_2O concentration of 100 ppm is clearly undercut. The optimization goal has been fulfilled and therefore it is possible to use the described thermal treatment system to reduce the N_2O concentration in the off-gas stream below the required boundary.

5.1.5 Conclusions for thermal treatment case study

The case study has highlighted the different steps and corresponding tools for the 'Modelling objective and system information'-work-flow, the 'Single-scale model construction'-work-flow, the 'Model identification'-work-flow as well as the 'Optimization'-work-flow (see Chapter 2). The parameter identification part (Phase III) has fine-tuned the parameter values derived from databases for the conditions of the experimental measurements. It has been shown that, applying methods like sensitivity and identifiability analysis can firstly reduce the size of the optimization problem significantly (from 157 to 9 to 5 parameters) and secondly to improve the quality of the model fit. The application of the model for the optimization problem has revealed that it is possible to apply the thermal treatment step to reduce the N_2O concentration in the

adipic acid off-gas stream below the required boundary of 100 ppm. The sensitivity analysis step within the optimization has further revealed that the design variable pressure has an opposite effect on the two optimization goals and that in general the impact of temperature is stronger than the impact of pressure.

5.2 Biotechnology industry: Batch protein uptake of Lysozyme by Sepharose beads

This case study involves the batch uptake of the protein lysozyme by sepharose beads. Models of different scales need to be developed so that the product behaviour (sepharose particles) can be monitored in a batch processing step. The case study has been chosen because through it the developed modelling features of the 'Multi-scale model construction'-work-flow (Figure 2.5) can be highlighted. The application of the multi-scale model construction work-flow to this case study has been published already (Heitzig et al., 2010), but in a slightly different manner. Large parts of this section however, are based on the publication. In the paper a top-down approach has been applied to develop the multi-scale model starting from a single-scale scenario. In this chapter a different approach is presented starting from two more complex multi-scale scenarios which are derived based on the modelling objective and the system information (Phase I). The third multi-scale scenario developed has not been published.

Like in any model development problem the starting point for this case study is the 'Modelling objective and system information'-work-flow (Figure 2.2).

5.2.1 Phase I. Modelling objective and system information

Step I.1: Modelling Objective (Phase I)

The modelling objectives here are to predict the change of the bulk concentration of the protein lysozyme originating from the protein uptake by sepharose beads in a batch (process) operation as well as the uptake profiles of the protein inside the sepharose beads. It is sufficient to consider the protein concentration in the bulk as lumped since the batch system is well-mixed. In order to be able to predict the protein uptake profiles the sepharose beads need to be discretized (at least) in radial direction.

Step I.2: System information and documentation (Phase I)

Step I.2.1: Functional description/ sketch of the system to be modelled (process, unit operation, product)

The system to be modelled is sketched in Figure 5.11. It consists of a well-mixed batch vessel which contains sepharose beads (solid phase) surrounded by an aqueous solution of the protein lysozyme and salts (liquid phase). The lysozyme is taken up by the sepharose particles.

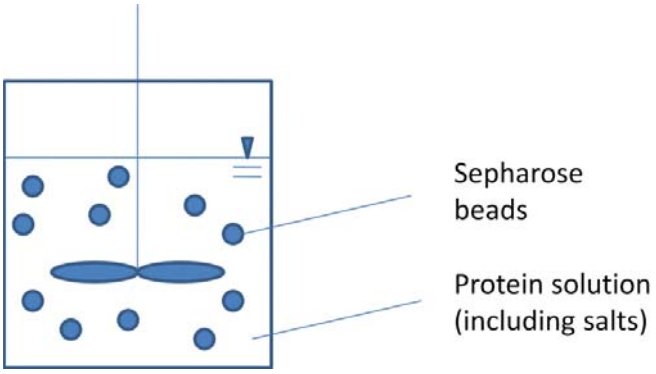


Figure 5.11 *Batch uptake of the protein lysozyme by sepharose beads.*

Step I.2.2 System conditions

For the batch process under consideration the salt concentration in the bulk solution is constant at 100 mM throughout operation. The temperature is ambient and the pressure at 1 atm. The mean diameter of the sepharose beads is 50 μm.

Step I.2.3 Phenomena in the system that might be of importance

Several phenomena that potentially can be of importance with respect to the modelling goal occur in the system:

- Convective mixing
- Mass transfer between bulk solution and sepharose beads
- Transport of protein inside particle pores (diffusion, pore diffusion)
- Adsorption of protein on sepharose surface

Step I.2.4 Modelling of system/ problem

For this case study all information on how the system can be modelled has been taken from different literature sources.

The protein uptake rate depends on the salt concentration of the protein solution. It increases at high salt concentrations (Dziennik et al., 2005). At the same time the final amount of adsorbed protein in equilibrium (given by the adsorption isotherm) decreases (Dziennik et al., 2005). The protein uptake curves, that is, the protein concentration fronts penetrating into the particles, are sharp at low salt concentrations and become more and more diffuse at higher salt concentrations (Dziennik et al., 2005). In general, however the transition from sharp to diffuse profiles for high salt concentrations does not occur for every adsorbent (Dziennik et al., 2003). Materials with a high surface charge density like the material under investigation here (sepharose) show this transition however (Dziennik et al., 2005). Since for the batch process under consideration the salt concentration is constant at 100 mM throughout operation the

model not necessarily needs to handle changing salt concentrations. The modelling objective formulated in Step I.1 needs to be extended by this information. If however the developed model is to be applied in a different application context, for example a chromatography process with salt gradient elution, the effects of changing salt concentration will become of importance and the modelling objective needs to be revised accordingly.

To model the uptake different approaches have been applied in literature all based on diffusion models (Dziennik et al., 2005). One diffusive model widely applied is pore diffusion (shrinking core model). Here, the protein transport is assumed to only occur in the pores whereas it is accompanied by adsorption of protein on the surface. This model however cannot be applied for high salt concentrations due to the fact that it cannot predict diffuse uptake profiles (Dziennik et al., 2005). For lysozyme and sepharose the model is applicable up to a salt concentration of 50 mM (Dziennik et al., 2005). An alternative is to model the uptake applying a homogenous diffusion model. The assumption here is that all protein in the sepharose particle (pores and surface) can diffuse and the driving force is given by the gradient of the total protein concentration in the pellet. The approach chosen for this case study considers parallel diffusion of pore and surface diffusion (Lenhoff, 2008) which can be applied for high and low salt concentrations. The surface diffusion is considered to cause the increase in the protein uptake rate with increasing salt concentrations.

Lenhoff et al. (2008) have developed a correlation for the ratio of apparent diffusion coefficient and pore diffusion coefficient with respect to changing salt concentrations and protein concentrations. Values for the correlation parameters are likewise provided.

Step I.2.5 Possible assumptions

Possible assumptions for the system are:

- Sepharose beads are spherical;
- Sepharose beads are of same size;
- Ideally mixed bulk solution;
- Transport in sepharose beads due to pore diffusion accompanied by adsorption on sepharose surface;
- Transport in sepharose beads due to parallel diffusion in pores and surface diffusion accompanied by adsorption of droplets on sepharose surface;
- Adsorption equilibrium is established inside the pellet at any time;
- Neglect mass transfer resistance between bulk solution and sepharose pores;
- Neglect impact of salt concentration on protein uptake characteristics (e.g. diffusion coefficients) ;
- Neglect impact of changing bulk protein concentration on protein uptake (adsorption and diffusion) ;

- Total protein uptake rate of sepharose beads small compared to protein bulk concentration.

Step I.2.6 Preliminary system/process/reactor data

Experimental data for model validation and/or identification are available from Dziennik et al. (2003, 2005):

- Measurements of protein concentration taken up of sepharose beads from bulk available from Dziennik et al. (2005) at a salt concentration of 100 mM and for high superficial velocities (2100 cm/h) (see Figure 5.16);
- Measurements of protein uptake profiles in sepharose beads with respect to bead radius from Dziennik et al. (2005) at a salt concentration of 100 mM and a superficial velocity of 500 cm/h (see Figure 5.17);
- Data for identification of adsorption isotherm parameters from Dziennik et al. (2003);
- Ratio of apparent diffusion coefficient and pore diffusion coefficient at a salt concentration of 100 mM: 3.25 (Dziennik et al., 2005).

From this follows that experimental data is already available to validate both modelling goals, the prediction of the protein bulk concentration and the simulation of the uptake profiles of the protein within the sepharose bead.

Step I.2.7 Select model-scenarios of interest

The degree-of-detail-determining factors reported in Chapter 2 support the modeller in this important work-flow step. Based on the modelling objective and the now available information on the system (possible phenomena, modelling of system in literature, available experimental data, possible assumptions) two different multi-scale scenarios are derived that are to be developed.

In order to fulfil the modelling goal a model-scenario needs to span at least two length scales. The highest length scale is the scale of the batch vessel. On this scale the protein bulk concentration can be monitored. In order to be able to predict the protein bulk concentration it might be required to increase the degree of detail and add smaller size scales. However, already the second part of the modelling objective demands for at least one smaller size scale which introduces a model for a single sepharose bead to the multi-scale scenario. This scale is required to be able to predict the protein uptake curves within the sepharose beads. The preliminary data search has furthermore revealed that experimental data to validate the multi-scale scenario on both of these two length scales is available. Consequently, the simplest scenario to be developed (scenario 1) considers the overall batch vessel on the macro scale and the sepharose particle on a micro scale. The apparent diffusivity is set to a constant experimental value taken from (Dziennik et al., 2005). However, the apparent diffusivity has been determined

to depend on and change with the protein and salt concentrations in the aqueous solution. Therefore, a second alternative multi-scale scenario is developed (scenario 2) which adds the correlation developed by Lenhoff et al. (2008) for the diffusion coefficient ratio (see Step 2.4) as a nano-scale model.

Based on the development and evaluation of the two identified initial multi-scale scenarios the optimal final scenario with respect to the modelling goal is derived during the iterative modelling procedure (Phases I-IV) by comparing the resulting scenarios and if necessary extending and/ or simplifying these scenarios.

5.2.2 Phase II. Model construction

The models for both multi-scale scenarios are constructed based on the information collected in the previous phase applying the multi-scale model construction work-flow (Phase II.B).

Step II.1: Model-scenario documentation and concept (Phase II)

First, scenario 1, the simplest reasonable multi-scale scenario with respect to the modelling objective, is developed. Two different scales are considered. Since the modelling objective demands for the prediction of the change of the protein bulk concentration the overall batch reactor is to be included in the multi-scale-scenario as the macro scale. Further, the uptake curves within the sepharose beads are to be predicted. This demands for at least inclusion of another scale (micro) in the scenario. The micro scale consists of the parallel diffusion model to determine the protein uptake rate and the concentration profiles within the sepharose beads. Figure 5.12 shows a schematic sketch of this first investigated scenario and its scales.

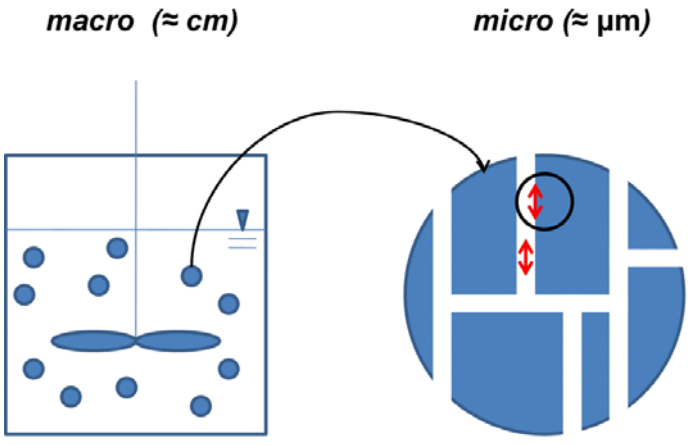


Figure 5.12 *Schematic sketch for scenario 1 (macro+micro scales)*

For scenario 2 the ratio of apparent and pore diffusion coefficient is not set to a constant experimental value but is replaced by a nano-scale model (Lenhoff, 2008) which depends on the salt and protein concentrations in the surrounding bulk solution. Figure 5.13 shows the corresponding sketch for scenario 2 and its scales.

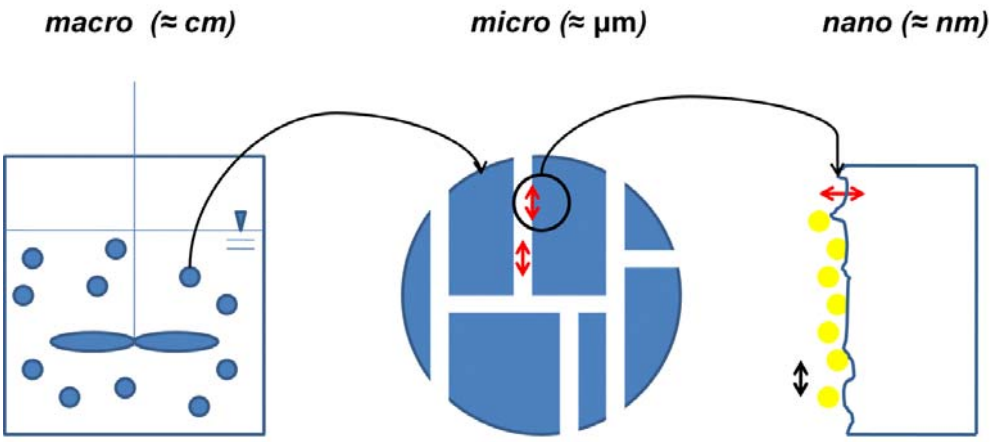


Figure 5.13 *Schematic sketch for scenario 2 (macro+micro+nano scales)*

The micro-scale describes, as shown in Figure 5.13, how fast the protein enters and penetrates the particle based on a diffusion model. The nano-scale model provides a deeper insight in the protein-wall interaction considering properties of the protein (lysozyme), the solid phase material (sepharose) and the protein adsorption energetics.

In practice, only one scenario is developed at a time. Here, the development of scenarios 1 and 2 is presented together in order to avoid repetition of the same text.

Assumptions valid for both multi-scale scenarios are:

- Ideally mixed bulk phase;
- micro scale model distributed in radial direction;
- sepharose beads of equal size and spherical shape;
- outer mass transfer resistance between bulk and sepharose beads negligible;
- inner protein transport within the sepharose beads can be described by a parallel diffusion model of surface and pore diffusion;
- adsorption equilibrium is established inside the pellet at all times;
- Langmuir isotherm can be applied for the adsorption equilibrium;
- pore diffusion coefficient is taken from literature for the diffusion of a protein in a pore without wall interaction (Carta et al., 2005);
- isothermal system;

For scenario 1 it is moreover assumed that the diffusion coefficients are constant. This means that the impact of changing protein concentrations on the diffusion coefficients is neglected and that the salt concentration remains constant during operation or that the influence of changing salt concentrations is negligible. For scenario 2 changing protein and salt concentrations are considered whereas it is assumed that the applied correlation for the prediction of the ratio of the apparent and pore diffusion coefficients (Lenhoff et al., 2008) delivers reliable predictions.
The considered phenomena for the two multi-scale scenarios are ideal convective mixing in the bulk solution, parallel diffusion of lysozyme within the sepharose beads and adsorption of lysozyme onto the sepharose beads surface. The integration framework (see Table 1.3) applied here is the multi-domain framework.

Step II.2: Model development for new/current scale (Phase II)

For each scale the relevant phases of the model development process (Phases I-IV) are repeated. The major part of the 'Modelling objective and documentation'-phase results from the information collected already for the overall scenario. However, it is of importance for the re-use of a scale-model in a different context to store the relevant information together with this model. Here, this phase is skipped however and it is directly continued with Phase II (the model construction). In order to construct the models for the different scales in the multi-scale scenario the single-scale model construction work-flow is applied in a loop over all scales currently in the multi-scale scenario.

Phase II.B Single-scale model construction

For each scale the model equations are derived (mostly from literature), introduced to ICAS-MoT, translated by the tool into a model object and analysed. In the following, the model equations for each scale are given. The numerical model analysis is described exemplary for the micro-scale model which is the most complex model in the multi-scale scenarios.

1. Macro scale model

The macro-scale (Figure 5.13, left) model equations are given below. Equation 5.9 determines the concentration of protein taken up by all particles while Equation 5.10 results the protein bulk concentration.

$$V_{particle} = \frac{4}{3}\pi R_{particle}^3 \tag{5.6}$$

$$V_{bulk} = V_{liq} + 0.5V_{suspension} \tag{5.7}$$

$$m_{ads} = \frac{0.5V_{suspension}}{V_{particle}} \cdot m_{ads}^P \tag{5.8}$$

$$q_{ads} = \frac{m_{ads}}{0.5V_{suspension}} \tag{5.9}$$

$$c_P^{bulk} = c_P^{bulk,start} - \frac{m_{ads}}{V_{bulk}} \tag{5.10}$$

Table 5.9 provides the values of the variables that need to be specified as known or parameters in order to satisfy the degree of freedom of the model.

Table 5.9 *Values of known variables for macro scale*

	values:
known variables:	
m_{ads}^P	calculated on micro scale
$c_P^{bulk,start}$	1.5 [mg/mL]
R	50 [µm]
$V_{suspension}$	1.6 [mL]
V_{liq}	79 [mL]

2. Micro scale model

For the micro scale (Figure 5.13, centre) model, the partial differential equation (PDE) describing the parallel diffusion into the sepharose bead has been taken from Lenhoff (2008) and discretised in radial direction during model translation using the method of lines resulting in a system of ordinary differential equations (ODEs). 50 discretisation points have turned out to be sufficient. The resulting model is given in Equations 5.11-5.267.

$$q_0 = \frac{q_{mon} \cdot b \cdot c_p^{bulk}}{1 + b \cdot c_p^{bulk}} \tag{5.11}$$

$$\alpha = \frac{q_0}{\varepsilon_p \cdot c_P^{bulk}} \tag{5.12}$$

$$\beta = \left(D_a D_p - 1\right)/0.6 \tag{5.13}$$

$$R_{sep} = \frac{1}{1 + b \cdot c_P^{bulk}} \tag{5.14}$$

$$\rho(0) = 0, \rho(50) = 1, \Delta\rho = \frac{\rho(50)}{50}, \rho(1) = 0.5 \cdot \Delta\rho \tag{5.15-5.18}$$

$$\rho(j) = \rho(j-1) + \Delta\rho, \quad j = 2, \dots, 49 \tag{5.19-5.66}$$

$$Y(50) = 1 \tag{5.67}$$

$$\frac{dY(1)}{\partial \tau_P} = \left\{ \frac{1}{2 \cdot \Delta\rho \cdot \rho(1)^2} \left(\frac{\rho(2)^2 \cdot R_{sep} \frac{Y(3) - Y(1)}{2\Delta\rho}}{\left(1 - \left(1 - R_{sep}\right) \cdot Y(2)\right)^2} - \frac{\rho(0)^2 \cdot R_{sep} \cdot \frac{Y(1) - \frac{4Y(2) - Y(3)}{3}}{\Delta\rho}}{\left(1 - \left(1 - R_{sep}\right) \frac{4Y(2) - Y(3)}{3}\right)^2} \right) + \frac{2\beta}{\rho(1)} \cdot \frac{Y(2) - \frac{4Y(2) - Y(3)}{3}}{2\Delta\rho} + \beta \cdot \frac{Y(2) - 2Y(1) + \frac{4Y(2) - Y(3)}{3}}{\Delta\rho^2} \right\} \cdot \left(\alpha + \frac{R_{sep}}{\left(1 - \left(1 - R_{sep}\right) Y(1)\right)^2} \right)^{-1} \tag{5.68}$$

$$\frac{dY(2)}{\partial\tau_P} = \left\{ \frac{1}{2\cdot\Delta\rho\cdot\rho(2)^2} \left(\frac{\rho(3)^2\cdot R_{sep}\cdot\frac{Y(4)-Y(2)}{2\Delta\rho}}{\left(1-\left(1-R_{sep}\right)\cdot Y(3)\right)^2} - \frac{\rho(1)^2\cdot R_{sep}\cdot\frac{Y(2)-\frac{4Y(2)-Y(3)}{3}}{2\Delta\rho}}{\left(1-\left(1-R_{sep}\right)\cdot Y(1)\right)^2} \right) + \frac{2\beta}{\rho(2)}\cdot\frac{Y(3)-Y(2)-1}{2\Delta\rho} + \beta\cdot \frac{Y(3)-2Y(2)+Y(1)}{\Delta\rho^2} \right\} \cdot \left(\alpha + \frac{R_{sep}}{\left(1-\left(1-R_{sep}\right)Y(2)\right)^2} \right)^{-1} \quad (5.69)$$

$$\frac{dY(j)}{\partial\tau_P} = \left\{ \frac{1}{2\cdot\Delta\rho\cdot\rho(j)^2} \left(\frac{\rho(j+1)^2\cdot R_{sep}\cdot\frac{Y(j+2)-Y(j)}{2\Delta\rho}}{\left(1-\left(1-R_{sep}\right)\cdot Y(j+1)\right)^2} - \frac{\rho(j-1)^2\cdot R_{sep}\cdot\frac{Y(j)-Y(j-1)}{2\Delta\rho}}{\left(1-\left(1-R_{sep}\right)\cdot Y(j-1)\right)^2} \right) + \frac{2\beta}{\rho(j)}\cdot\frac{Y(j+1)-Y(j)-1}{2\Delta\rho} + \beta\cdot \frac{Y(j+1)-2Y(j)+Y(j-1)}{\Delta\rho^2} \right\} \cdot \left(\alpha + \frac{R_{sep}}{\left(1-\left(1-R_{sep}\right)Y(j)\right)^2} \right)^{-1}, j = 3-48 \quad (5.70\text{-}5.115)$$

$$\frac{dY(49)}{\partial\tau_P} = \left\{ \frac{1}{2\cdot\Delta\rho\cdot\rho(49)^2} \left(\frac{\rho(50)^2\cdot R_{sep}\cdot\frac{Y(50)-Y(49)}{\Delta\rho}}{\left(1-\left(1-R_{sep}\right)\cdot Y(50)\right)^2} - \frac{\rho(48)^2\cdot R_{sep}\cdot\frac{Y(49)-Y(48)}{2\Delta\rho}}{\left(1-\left(1-R_{sep}\right)\cdot Y(48)\right)^2} \right) + \frac{2\beta}{\rho(49)}\cdot\frac{Y(50)-Y(49)-1}{2\Delta\rho} + \beta\cdot \frac{Y(50)-2Y(49)+Y(48)}{\Delta\rho^2} \right\} \cdot \left(\alpha + \frac{R_{sep}}{\left(1-\left(1-R_{sep}\right)Y(49)\right)^2} \right)^{-1} \quad (5.116)$$

$$q(j) = Y(j)\cdot q_0, \quad j = 1,\dots,50 \quad (5.117\text{-}5.166)$$

$$c(j) = \frac{q(j)}{b\cdot(q_{mon}-q(j))}, \quad j = 1,\dots,50 \quad (5.167\text{-}5.216)$$

$$V_{cell}(j) = \frac{4\pi}{3}\cdot\left(\left((\rho(j)-0.5\Delta\rho)\cdot R + \Delta\rho\cdot R\right)^3 - \left((\rho(j)-0.5\Delta\rho)\cdot R\right)^3\right), j = 1,\dots,50 \quad (5.217\text{-}5.266)$$

$$m_{ads}^P = \sum_{j=1}^{50}(q(j)\cdot V_{cell}(j)) + \sum_{j=1}^{50}\left(c(j)\cdot V_{cell}(j)\cdot\varepsilon_p\right) \quad (5.267)$$

The boundary conditions of the PDE for parallel diffusion are given below:

$$Y(50) = 1 \quad (5.268)$$

$$\frac{dY}{d\rho}(0) = 0 \quad (5.269)$$

Equation 5.259 is not part of the discretized model because the boundary condition has been implicitly included in the model equations 5.11-5.267.

After model translation, a numerical model analysis is conducted. In a first step, the equations are classified. The micro-scale model contains 208 algebraic equations (AEs) and 49 ordinary differential equations (ODEs). In a next step, the model variables are pre-classified in dependent, independent and algebraic variables. Table 5.10 summarizes the results.

Table 5.10 *Pre-classification of variables (micro-scale)*

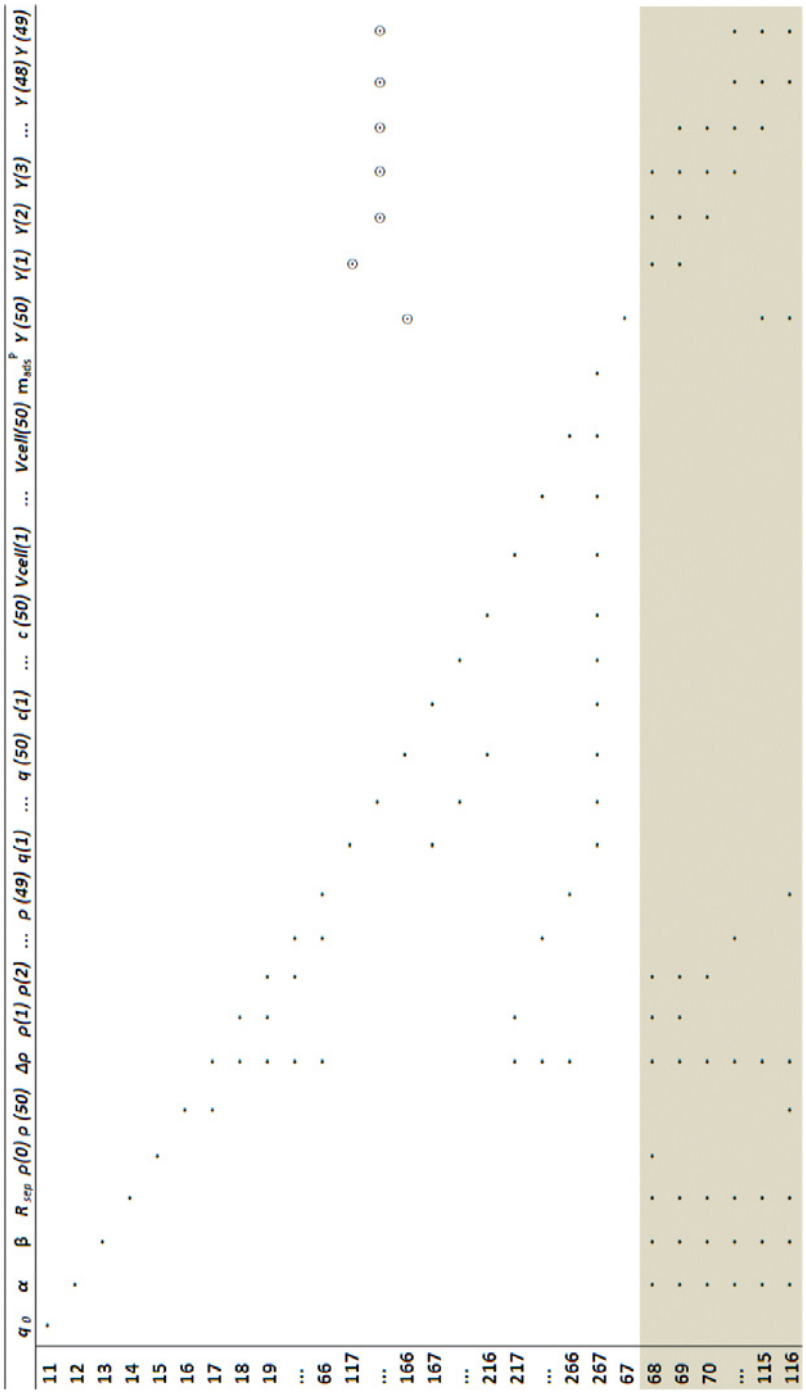
type	***number***	***variable names***
independent	1	τ_p
dependent	49	Y(j), j=1,…,49
algebraic variables	214	q_0, q_{mon}, b, c_p^{bulk}, α, β, ε_p, D_aD_p, R_{sep}, $\Delta\rho$, R, $Y(50)$, m_{ads}^P $\rho(j)$, $q(j)$, $c(j)$, $V_{cell}(j)$, j=1,…,50

The degrees of freedom are determined and satisfied separately first for the AE-system and then for the ODE-system. The degrees of freedom for the AE-part, which is calculated by the difference of the number of variables appearing in the AEs and the number of AEs, result to be 6 (214 variables-208 AEs). Consequently, 6 variables that appear in the AEs need to be specified either as known or parameter. The parameters of the Langmuir isotherm b and q_{mon} are specified as parameters. They are assigned an initial estimate and will be estimated with absorption isotherm data in a later step. Furthermore, ε_p, $D_a D_p$, R and c_p^{bulk} are specified as known. The degrees of freedom for the ODE part are given by the number of variables which are unknown and do not appear in the AEs. In this case it turns out to be 0 and no further variables need to be specified.

The incidence matrix of the system with optimized equation ordering is given in Table 5.11.

The matrix is divided in the algebraic equation part (upper part, white) and the ordinary differential equations part (lower part, dark). It can be seen that the algebraic equations are all explicit and have a lower tridiagonal form. The AE-part is however coupled with the ODE-part (off-diagonal elements). The ODEs are likewise coupled because the ODE-part of the incidence matrix cannot be transformed to a lower tridiagonal form.

Table 5.11 *Incidence matrix for micro-scale model*



Finally, the values for the known variables and parameters need to be given. For the two Langmuir parameters an initial estimate is assigned. Table 5.12 list the given and assumed values for the 6 variables.

Table 5.12 *Values of known variables and initial estimates for parameters for micro scale*

	values:
known variables:	
c_P^{bulk}	initial value: 1.5 [mg/mL] (from macro scale)
$D_a D_p$	3.25 [-] (Dziennik et al., 2005), calculated on nano scale for scenario 2
ε_p	0.71 [-] (DePhillips & Lenhoff, 2000)
R	50 [µm]
parameters:	
q_{mon}	1 [mg/mL] (initial estimate)
b	1 [mL/mg] (initial estimate)

Since the model contains ODEs initial conditions need to be given. It is assumed that the stationary phase does not contain any lysozyme at the beginning of the batch process.

$$Y(j) = 0 \; for \; \tau_P = 0, \quad j = 1, \ldots, 49 \tag{5.270}$$

For multi-scale scenario 1 the micro and macro scale models are combined. For scenario 2 the same models are needed, additionally the nano scale model is added (Figure 5.13, right).

3. Nano scale model

The nano scale model equations are given below (Lenhoff, 2008):

$$k_{retention} = 0.0017 \cdot IS^{-5.75} \tag{5.271}$$

$$\kappa = 3.29 \cdot IS^{0.5} \tag{5.272}$$

$$q_0 = \frac{q_{mon} \cdot b \cdot c_P^{bulk}}{1 + b \cdot c_P^{bulk}} \tag{5.273}$$

$$D_a D_p = 1 + 0.35 \cdot \left(\frac{q_0}{\varepsilon_p \cdot c_P^{bulk} \cdot \phi_p \cdot h} \right) \cdot \left(\frac{\phi_c}{k_{retention} \cdot \kappa} \right)^{0.56} \tag{5.274}$$

Table 5.13 lists the values of the known variables and the initial values of the parameters.

Table 5.13 *Values of known variables and initial estimates for parameters for nano scale*

	values:
known variables:	
c_P^{bulk}	initial value: 1.5 [mg/mL] (from macro scale)
IS	100 [mM]
h	3.5 [nm] (Lenhoff, 2008)
ε_p	0.71 [-] (DePhillips & Lenhoff, 2000)
ϕ_p	0.067 [nm^{-1}] (DePhillips & Lenhoff, 2000)
ϕ_c	0.053 [nm^{-1}] (DePhillips & Lenhoff, 2000)
parameters:	
q_{mon}	1 [mg/mL] (initial estimate)
b	1 [mL/mg] (initial estimate)

Phase III. Model identification

Model identification is needed for the micro-scale model. The two parameters of the Langmuir isotherm are fitted to experimental data. This is a very simple parameter estimation problem and steps like sensitivity and identifiability analysis are not considered. Instead, the parameter estimation step (Step 3) is performed directly. Experimental measurements of the isotherm for the adsorption of lysozyme on sepharose at a salt concentration of 100 mM are available from Dziennik et al. (2003) (22 datapoints). The initial values for the parameters have been set to 1, the lower bounds to 0 and the upper bounds to 2000. A least square fit objective function has been applied. The problem converges after 55 iterations (SQP optimizer in MoT) and the results are summarized in Table 5.14.

Table 5.14 *Results from parameter estimation for Langmuir isotherm*

q_{mon}	b	***OBJ***
122.98 mg/mL	21.49 mL/mg	1.497

Step II.3: Establish data-flow scheme for new current scale (Phase II)

The data-flow scheme has to be updated each time one of the models in a multi-scale scenario has been developed. Here, only the final data-flow schemes for multi-scale scenarios 1 and 2 are shown (Figure 5.14).

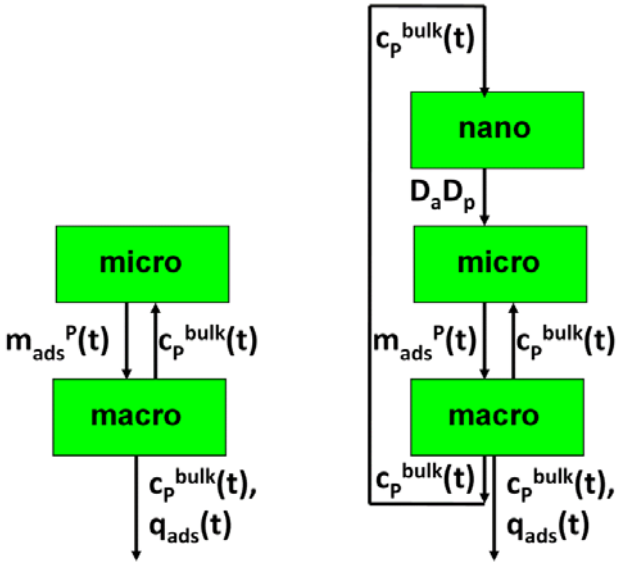


Figure 5.14 *Final data-flow schemes for multi-scale scenarios 1 (left) and 2 (right).*

A multi-scale analysis is not performed because at this point it is already decided which multi-scale scenarios are to be developed. If the performance of these scenarios turns out to be not satisfactory one option would be to come back to this step and identify potentials to increase the degree of detail which also might add new scales to the scenario.

Step II.4: Derive linking scheme and link models accordingly (Phase II)

Figure 5.15 shows the linking schemes of multi-scale scenario 1 and 2.

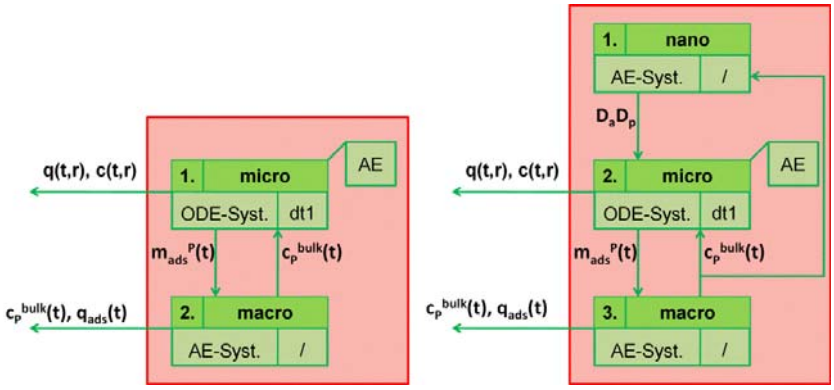


Figure 5.15 *Linking schemes multi-scale scenario 1 (left) and 2 (right).*

The small boxes in Figure 5.15 represent the different scales and give their types of equations as well as their time-scales. It can be seen that for both scenarios only one scale with differential

equations exists (micro scale: time-scale dt1). Consequently, the time-scale issue is not of importance here. For that reason all equations can be written to one overall model file for each scenario. The linking schemes further show the data-flow between the scales as well as the model output variables from each scale. The numbers in the boxes represent the solution sequence. The large box that surrounds both scales indicates that they need to be solved in a coupled manner which is obvious from the data-flow. For each time step of the ODE-scale the AE-scales need to be solved to convergence (nested, but here: AEs are all explicit).

Step II.5: Overall model analysis (Phase II)

In this step a numerical model analysis for the linked multi-scale model is made. This ensures the consistency of the different scale-models. The analysis of the incidence matrix can be used to confirm/adjust the optimal sequence of equations as well as for the identification of model parts that need to be solved coupled which have already been derived based on the data-flow between the scales in the previous work-flow step. The steps of the numerical analysis are analogous to the single-scale model construction work-flow and have been presented in detail for the micro scale model, and therefore, not described here.

5.2.3 Phase IV. Model evaluation/ validation

The different multi-scale scenarios are compared with respect to the fulfilment of the modelling objective. For this case study the modelling objective consists of two main parts, the first part is related to the model predictions on the macro-scale while the second to the nano-scale model predictions.

1. Modelling objective part I (macro scale)

For the macro scale the model predictions are the same for both multi-scale scenarios. Figure 5.16 shows the validation of the macro scale predictions. It compares the model predictions (solid line) to the experimental data (dotted line) for the amount of protein taken up by the sepharose versus time. The experimental data is available from Dziennik et al. (2005) at a salt concentration of 100 mM and for high superficial velocities (2100 cm/h).
The model predictions are satisfactory on the macro scale.

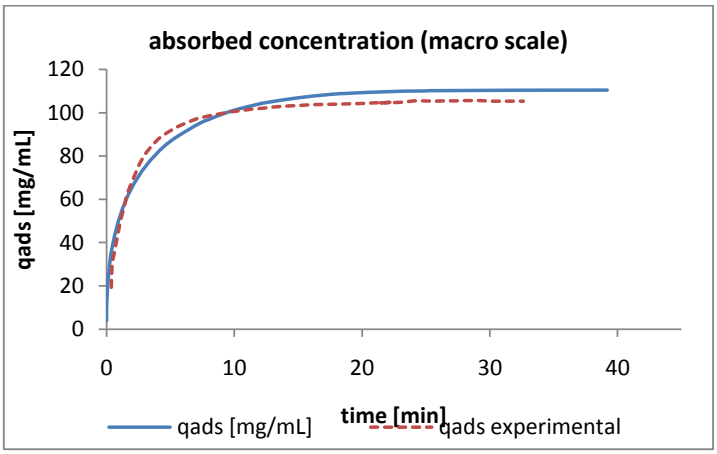


Figure 5.16 *Absorbed protein concentration by sepharose beads vs. time (simulated and experimental) for scenarios 1 and 2.*

2. Modelling objective Part II (nano scale):

The validation of the predicted uptake profiles within a sepharose bead is shown in Figure 5.17. The experimental data is available from Dziennik et al. (2005) who measured the profiles by confocal microscopy at a salt concentration of 100 mM and a superficial velocity of 500 cm/h.

First, scenario 1 is evaluated. It can be seen that there are deviations between the experimental measurements of the protein uptake profiles and the simulations. However, the profiles for 11 and 26 minutes are qualitatively good. The profiles for t=2 and t=5 minutes deviate a lot more, especially at the particle surface. This has also been stated by Dziennik (2005) where the system has been modelled applying a homogeneous diffusion model. The authors suggest the deviations to be caused by with the neglect of the external mass transfer resistance, which plays an important role at the low superficial velocity of 500 cm/h applied during the experimental measurements. The performance of the more complex scenario 2 turns out to be worse than the performance of scenario 1. This is due to the fact that the applied correlation for $D_a D_p$ on the nano scale introduces additional uncertainties to the scenario. For a salt concentration of 100 mM and a lysozyme bulk concentration of 1 mg/mL the measured value for $D_a D_p$ is about 3.3 (Dziennik et al., 2005) whereas the predicted value (with scenario 2) is around 2. From this can be concluded that due to the quality of the predictions of the nano-scale model for this case study, the less complex scenario 1 in which the nano scale is replaced by a constant experimental value provides the better results. This shows that the more complex scenario does not automatically give the most accurate results and that the predictions depend very much on the quality of the models for the different scales. In any case, a more complex scenario always increases the flexibility with respect to changing conditions. For example, for this case study, scenario 2 can handle the impact of changing protein bulk concentrations and ionic strength on the apparent diffusion coefficient. So, for applications where the salt and/or protein

concentrations change heavily during the process scenario 2 might be advantageous. This shows how much the derivation of the best-suited scenario depends on the application context and the modelling objective.

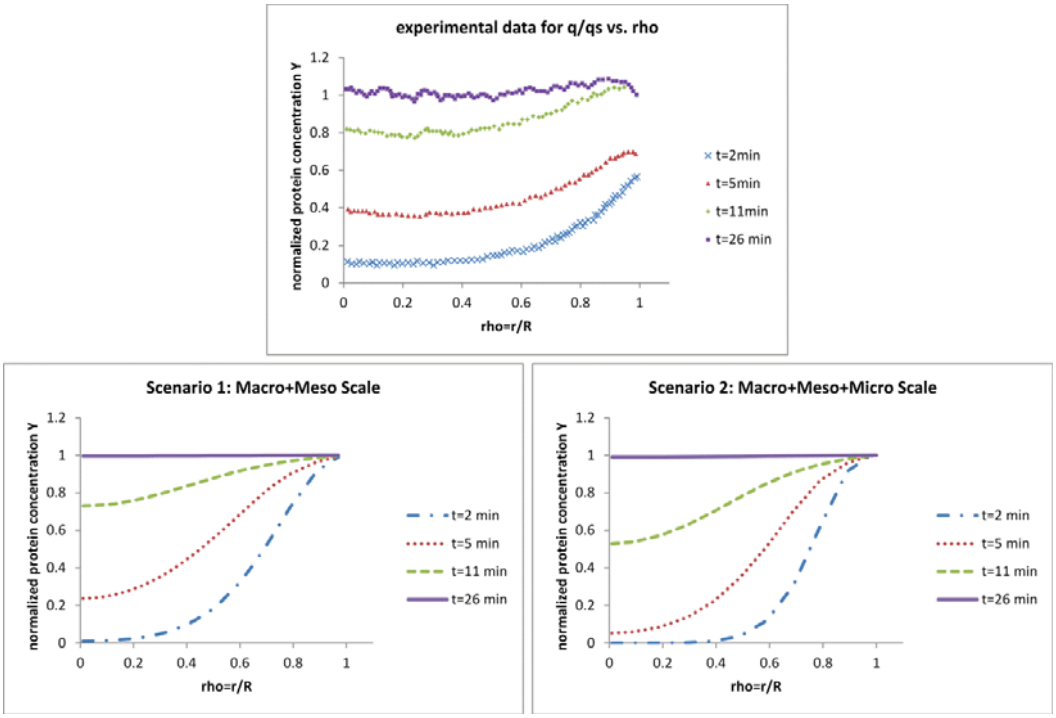


Figure 5.17 *Uptake profile of lysozyme in sepharose particle for time=2 min, 5 min, 11 min, 26 min, respectively; experiment (top), scenario 1 (left) and scenario 2 (right).*

Scenario 1 is selected to be the best scenario for the studied application context. However, the performance of scenario 1 is not satisfying and ways to improve the performance of the multi-scale scenario need to be investigated. In Chapter 2 a number of options to improve a model-scenario are provided. One option is to go back to the documented knowledge about the system in Phase I and based on that, increase or decrease the degree of detail by, for example, including an additional phenomenon that might be of importance. For this case study, it has been suggested during Phase I that the outer mass transfer resistance might play an important role and it has become obvious during model evaluation and validation that it cannot be neglected for low superficial velocities. Consequently, it has been decided to develop a third scenario (scenario 3) based on scenario 1 which considers the outer mass transfer resistance. In order to do so, the modeller needs to go to the multi-scale model construction work-flow and adjust the assumptions accordingly in Step II.1. Afterwards, the rest of the work-flow is followed to develop the new scenario (not shown here). The macro-scale model remains unchanged and for

the micro-scale model the boundary condition Y(50)=1 is replaced by the following differential equation:

$$\frac{dY50}{d\tau_p} = k \cdot (1 - Y50)/q_0 \tag{5.275}$$

Here, k is the outer mass transfer resistance coefficient. The value of k for a salt concentration of 100 mM and a superficial velocity of 500 cm/h has been determined by parameter estimation using the experimental data shown in Figure 5.17. The resulting value for k is 69.72 mg/(mL sec). Figure 5.18 shows the uptake profiles obtained for scenario 3.

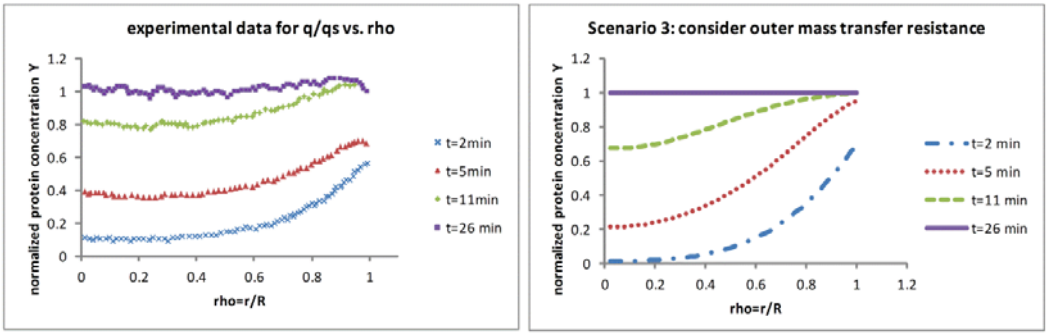


Figure 5.18 *Uptake profile of lysozyme in sepharose particle for time=2 min, 5 min, 11 min, 26 min, respectively; experiment (left) and scenario 3 (right).*

It can be seen that scenario 3 predicts the profiles by far better than the previous scenarios (especially close to the particle surface) and therefore is selected as the final optimal scenario.

5.2.4 Conclusions for protein uptake case study

The protein uptake case study has highlighted the application of the 'Modelling objective and system information'-work-flow and the 'Multi-scale model construction'-work-flow introduced in Chapter 2.

Two alternative starting multi-scale scenarios (scenarios 1 and 2) have been derived based on the modelling objective and the collected system information in Phase I. A third scenario (scenario 3) has been derived based on the evaluation of scenarios 1 and 2.

The model evaluation and validation in Section 5.2.3 has revealed that scenario 3 is the best option with respect to the modelling objective and that the mass transfer resistance between bulk and sepharose beads cannot be neglected if the experimentally determined profiles close to the particle surface are to be matched with the model. It needs to be mentioned in this context that the transfer coefficient has been estimated applying the protein uptake profiles, which were later also used for validation – this was not done for scenarios 1 and 2.

For the case that the model has to be extrapolable to different salt concentrations (e.g. application of the model in a chromatography process with salt gradient elution) scenario 2 is preferred and has to be extended to consider the outer mass transfer resistance. It might be advisable to improve the model predictions for the nano scale of scenario 2. This can be done by finding an improved expression for $D_a D_p$. One way to achieve this is the introduction of an adjustable parameter since the expression developed by Lenhoff et al. (2008) is purely predictive (not fitted to any experimental data). In general, the advantage of purely predictive models is their extrapolability to other systems (here, e.g. beads and/or proteins).

5.3 Chemical industry: Fluidized bed reactor

This case study demonstrates the construction and simulation of a multi-scale model-scenario for a fluidized bed reactor. The multi-scale scenario applied is adapted from Luss & Amundson (1968). It is demonstrated how the eigenvalue analysis reveals the different time-scales within the scenario and how it also indicates a potential for model simplification. Based on the results from the eigenvalue analysis a simplified model-scenario which only includes one scale is derived. This single-scale scenario has also been presented by Luss & Amundson (1968). However, these authors have not derived the single-scale scenario based on an eigenvalue analysis. Like the previous case study (Section 5.2) also this case study demonstrates the application of the multi-scale model construction work-flow. Only for this case study the application of the top-down approach is demonstrated to systematically increase the degree of detail starting from a simple single-scale scenario. Application of the simulation work-flow is also highlighted in this case study.

5.3.1 Phases I-III. Systematic derivation of two alternative model scenarios applying top-down strategy

Step I.1: Modelling objective (Phase I)

A model for the fluidized bed reactor in order to predict the reactant partial pressures in the fluidized bed and at the reactor outlet with respect to time needs to be developed.

Step I.2: System information and documentation (Phase I)

Step I.2.1 Functional description/sketch of system to be modelled

A sketch of the fluidized bed reactor to be modelled in this case study is given in Figure 5.19.
The reactor contains a fluidized bed of solid catalyst particles surrounded by a turbulent gas flow of the reactants/ products mixture. The reactants enter at the bottom of the reactor passing through a distributor. Turbulence is generated by the distributor and the high velocity of the gas stream. The distributor also prevents catalyst particles from leaving the reactor at the bottom. The turbulent gas stream stirs and fluidizes the catalyst bed which results in a uniform particle mixing and temperature distribution in the bed. The product stream exits the reactor at the top and is passed through a cyclone in order to remove solid catalyst particles (Luss & Amundson, 1968).

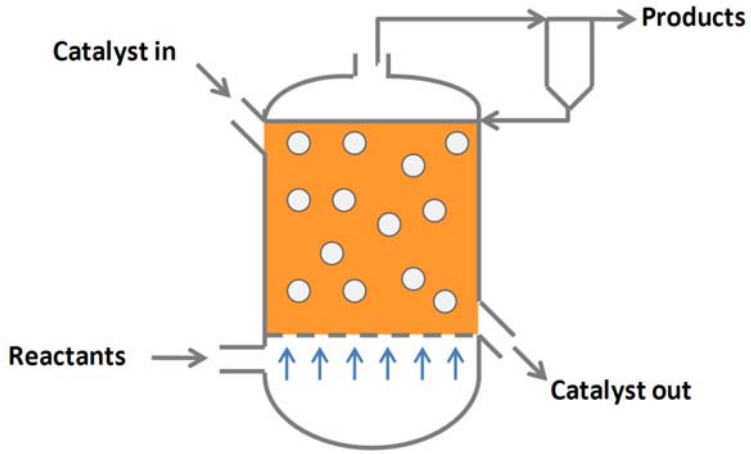


Figure 5.19 *Sketch of fluidized bed reactor.*

Step I.2.2 System conditions

The pressure of the system is atmospheric (1 atm). The walls of the fluidized bed reactor contain heating elements and have a constant temperature of 720 R. The partial pressure of the reactant in the input stream to the reactor is 0.1 atm and its temperature is 600 R. These values have been chosen according to Luss & Amundson (1968).

Step I.2.3 Phenomena in system that might be of importance

Phenomena that might play a role in the system under consideration are:

- Gas phase reaction;
- Formation of catalyst particles;
- Contact between particles and reactants in gas phase;
- Turbulent flow field;
- Transport of particles in flow field;
- Transport of fluid in gas phase (dispersion, diffusion, convection);
- Transport of fluid from bulk to catalyst surface (diffusion)/ mass transfer resistance between bulk and pellets;
- Transport of fluid inside particles: in pores, on particle/pore surface (e.g. diffusion);
- Chemical adsorption of the reactants on the catalyst surface;
- Heat transport in fluid phase (dispersion, heat conduction, convection);
- Heat transfer between fluid phase and fluidized catalyst bed/ heat transfer resistance;
- Heat transport in particle (conduction, convection, dispersion);
- Heat of reaction;
- Heat transfer between walls and gas phase;

- Radiation;
- Different temperatures of the catalyst pellets;
- Varying size of catalyst pellets;
- Cyclone;

Step I.2.4 Modelling of system/ problem

Luss & Amundson (1968) have modelled the fluidized bed reactor considering two different size scales. The macro scale model reflects the overall reactor and the micro scale a single catalyst pellet. On both scales a system of ODEs needs to be solved. The macro-scale mass and energy balances result the partial pressures of the reactants and the temperature in the gas phase. These are communicated to the micro-scale model which calculates and communicates back the partial pressures of the reactants and the temperature in the catalyst pellet. The catalyst pellets are assumed to be of equal size distribution. The gas is assumed to be ideal mixed and likewise inside the pellets the temperature and partial pressures are assumed to be the same everywhere. The authors have also suggested a simplification of the described model by neglecting the heat and mass transfer resistance between the gas phase and the catalyst pellets. This simplification leads to a single scale model.

Step I.2.5 Possible assumptions

Possible assumptions for the fluidized bed reactor model are (Luss & Amundson, 1968):

- Reactants in the gas phase are ideally mixed throughout the whole bed inside the reactor;
- Changes in the void fraction volume of the bed due to reaction are neglected;
- Particles are assumed to be small enough to consider that heat and mass transfer resistances can be lumped at the particle surface;
- Heat and mass transfer resistances at particle surface can be neglected;
- Reactions take place in the porous volume of the catalyst;
- Neglect existence of catalyst particles, reaction takes place in gas bulk;
- All particles have the same spherical shape and size;
- The initial temperature is the same for all particles in the bed;
- All particles have the same partial pressure of the reactants;
- One irreversible reaction is considered *A*->*B*;
- The amount of particles is constant during the whole operation;
- Cyclone operation found in the top of the reactor is not taken into account in the mathematical model.

Step I.2.6 Preliminary process/ system/ reactor data

No preliminary experimental data has been found.

Step I.2.7 Select model-scenarios of interest

If the modelling objective is considered it does not necessarily demand for a multi-scale model as in the case for the protein uptake case study (in Section 5.2). The only scale definitely required with respect to the modelling objective is the overall reactor scale. For that reason it is decided to start with a model-scenario which only considers the overall reactor scale and apply the top-down strategy to increase the degree of detail if necessary to improve the model predictions. Since there is no experimental data for validation, the starting single-scale scenario is compared to the more complex multi-scale scenarios resulting from the top-down strategy. The multi-scale model construction work-flow is applied in order to construct the starting model scenario and based on this the more complex multi-scale scenario is derived.

Step II.1: Model scenario documentation and concept (Phase II)

The first model scenario (scenario 1) only considers the starting scale, that is, the overall reactor. This infers that the existence of the catalyst particles is neglected and the reaction is assumed to take place in the fluid bulk phase. In addition, the following assumptions are made for this model-scenario:

- Reactants and products in the gas phase are ideally mixed throughout the whole bed;
- Changes in the volume of the bed due to reaction are neglected;
- One irreversible reaction is considered *A*->*B*;
- Reaction rate and heat can be represented applying the Arrhenius approach;
- Product of molecular weight and total pressure is constant throughout the whole bed;

The considered phenomena for scenario 1 can be listed as:

- Reversible reaction A->B in bulk gas phase;
- Convective mixing for fluid phase (turbulence);
- Heat of reaction;
- Heat transfer between walls and gas phase.

Step II.2: Model development for new/current scale (Phase II)

The only scale in the model-scenario is the reactor scale. The corresponding model is developed following the relevant work-flows of the model development process (Phases I-IV). The major part for the model documentation in Phase I can be copied from the previous step. Afterwards, the single-scale model construction work-flow is applied. For each scale (here only overall reactor scale) the model equations are derived, introduced to ICAS-MoT, translated into a model

object by the tool and analysed. This is not shown in detail here. Only the model equations are provided together with the main results from the numerical model analysis.

The model equations for scenario 1 have been derived from the 2-scale model given by Luss & Amundson (1968).

$$k_k = 0.0006 \cdot exp\left\{\frac{20.7-15000}{T}\right\} \tag{5.276}$$

$$\frac{dp}{d\tau^*} = p_e - p - \mathrm{H}_\mathrm{g}^* \cdot \mathrm{k} \cdot p \tag{5.277}$$

$$\frac{dT}{d\tau^*} = T_e - T + H_w(T_w - T) + \mathrm{H}_\mathrm{T}^* \cdot (-\Delta H) \cdot \mathrm{k} \cdot p \tag{5.278}$$

Here, p and T are the partial pressure and temperature of the reactant in the reactor, p_e and T_e are the pressure and the temperature at the reactor entrance. T_w is the temperature of the reactor wall. Further, k is the rate of the catalysed reaction and ΔH the heat of reaction. τ^* is the dimensionless time. H_g^*, H_w and H_T^* are coefficients given in Equations 5.280-5.282.

$$\tau^* = \frac{q \cdot t}{V \cdot \rho_g} \tag{5.279}$$

$$H_g^* = \frac{MW \cdot P \cdot V}{q} \tag{5.280}$$

$$H_w = \frac{2 \cdot h_w \cdot V}{r \cdot c_g \cdot q} \tag{5.281}$$

$$H_T^* = \frac{V}{q \cdot c_g} \tag{5.282}$$

In the above equations q is the gas mass flow rate, t is the time, V is the bed volume, ρ_g is the gas density, MW is the molecular weight, P is the total pressure, h_w is the heat transfer coefficient between the reactor wall and the gas, r is the radius of the fluidized bed and c_g is the heat capacity of the gas.

The model consists of 1 algebraic equation (AE) and 2 ordinary differential equations (ODEs). The variables can be pre-classified as 1 independent variable (τ^*), 2 dependent variables (p, T) and 8 algebraic variables (p_e, T_e T_w, k, H_g^*, H_w, H_T^*, ΔH). From the algebraic variables 7 variables need to be specified in order to satisfy the degree of freedom. All algebraic variables are specified as known except for k which is an explicit algebraic variable. Table 5.15 summarizes the variable values assigned to the known variables as well as the initial conditions for the dependent variables.

Table 5.15 *Variable values for scenario 1 (values taken from Luss & Amundson, 1968)*

Dependent	p	0.1 [atm]
	T	600 [°R]
Known	p_e	0.1 [atm]
	T_e	600 [°R]
	T_w	720 [°R]
	H_g^*	320 $\left[\frac{atm \cdot hr \cdot ft^3}{lb.mol}\right]$
	H_w	1.6 [-]
	H_T^*	0.053333 $\left[\frac{F \cdot hr \cdot ft^3}{B.t.u}\right]$
	ΔH	-8x10^4 [B.t.u./lb. mol]

The incidence matrix is shown in Table 5.16.

Table 5.16 *Incidence matrix for scenario 1:*

	k	p	T	
(5.276)	*		⊛	AE-part
(5.277)	*	*		ODE-part
(5.278)	*		*	

The incidence matrix shows that the ODE-part is of lower tridiagonal from but that the algebraic equation and the ODE-part are coupled and cannot be solved separately (see off-diagonal element). Consequently, in each time step, the algebraic equation needs to be evaluated for the current values of the independent variables to obtain updated values for the RHS (right hand side) of the ODEs. In ICAS-MoT the BDF solver is chosen which can handle stiff and non-stiff dynamic systems.

Step II.3: Establish data-flow scheme for new/current scale (Phase II)

Figure 5.20 shows the data-flow scheme for scenario 1.

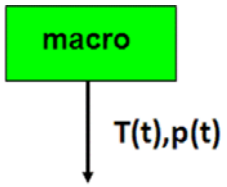


Figure 5.20 *Data-flow scheme for fluidized bed reactor scenario 1.*

Step II.4: Derive linking scheme and link models accordingly (Phase II)

This step is not needed because scenario 1 only contains one scale.

Step II.5: Overall model analysis (Phase II)

Since the model-scenario contains only one scale the overall model analysis is identical to the above described numerical model analysis for the model of the overall reactor scale.

A parameter estimation is not needed. Consequently, the next step is to evaluate the developed scenario. Until now it cannot be evaluated if the model performance is satisfactory due to the lack of data. However, the degree of detail of the current scenario should be increased in order to see how the performance of the scenario changes. For the new scenario (scenario 2) an additional scale for the catalyst particles is added. Scenario 2 is identical to the model proposed by Luss & Amundson (1968). The new scenario has been derived based on the information collected in Phase I and the modeller needs to go through Phases II-IV applying the multi-scale model construction work-flow in Phase II in order to develop the scenario. This is not shown in detail here but instead the main model features (model equations, linking scheme, overall model analysis results) are highlighted in the following sub-section.

Alternative multi-scale scenario with increased degree of detail (scenario 2)

Scenario 2 consists of two scales and has been taken from Luss & Amundson (1968). The macro scale considers the overall fluidized bed whereas the micro scale models the catalyst particle. The model equations for the macro scale are given below:

$$\frac{dp}{d\tau} = p_e - p + H_g \cdot (p_p - p) \tag{5.283}$$

$$\frac{dT}{d\tau} = T_e - T + H_w \cdot (T_w - T) + H_T \cdot (T_p - T) \tag{5.284}$$

Here, p and p_p are the partial pressures of the reactant in the steam and inside the catalyst particle, respectively. T and T_p are the corresponding temperatures. For scenario 2 the equations for the dimensionless time (τ) and factors (H_g, H_w and H_T) differ from scenario 1:

$$\tau = \frac{q \cdot t}{\varepsilon \cdot \rho_g V} \tag{5.285}$$

$$H_g = \frac{a_v \cdot k_g \cdot M \cdot P \cdot V}{q} \tag{5.286}$$

$$H_w = \frac{2 \cdot h_w \cdot V}{r \cdot c_g \cdot q} \tag{5.287}$$

$$H_T = \frac{a_v \cdot h_g \cdot V}{q \cdot c_g} \tag{5.288}$$

Here, ε is the void fraction of the bed, a_v is the interfacial area between the catalyst and the bulk fluid per unit volume, k_g is the mass transfer coefficient between the bulk fluid and the particles and h_g is the heat transfer coefficient between the gas and the catalyst particles. The model equations for the micro scale are given by:

$$\mathrm{k} = 0.0006 \cdot exp\left\{\frac{20.7-15000}{T_p}\right\} \tag{5.289}$$

$$\frac{dp_p}{d\tau} = \frac{H_g}{A} \cdot \left(p - p_p\right) - \frac{H_g \cdot K_k \cdot p_p}{A} \tag{5.290}$$

$$\frac{dT_p}{d\tau} = \frac{H_t}{C} \cdot \left(T - T_p\right) + \frac{H_T \cdot F \cdot K_k \cdot p_p}{C} \tag{5.291}$$

The factors A, C and F are given as:

$$A = \frac{\alpha \cdot v_p \cdot a_v}{\varepsilon \cdot s_p} \tag{5.292}$$

$$C = \frac{a_v \cdot c_s \cdot v_p \cdot \rho_s}{\varepsilon \cdot s_p \cdot c_g \cdot \rho_g} \tag{5.293}$$

$$F = \frac{(-\Delta H) \cdot k_g}{h_g} \tag{5.294}$$

Here, α is the void fraction of the particles, v_p is the catalyst particle volume, s_p is the particle surface and c_s is the heat capacity of the particles. Figure 5.21 shows the linking scheme for scenario 2.

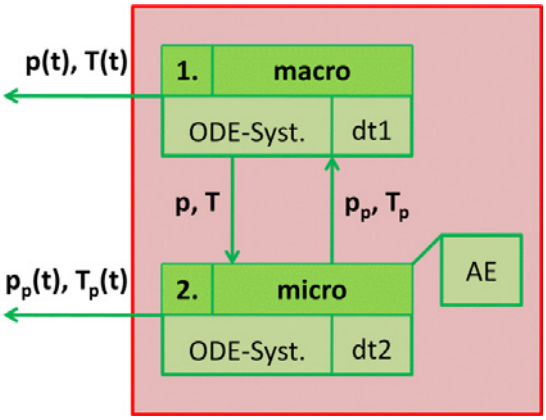


Figure 5.21 *Linking scheme for fluidized bed reactor scenario 2.*

The linking scheme shows the different scales, their output variables, time-scales (dt1, dt2), equation types and the data-flow between the scales. The big box in Figure 5.21 symbolizes that the scales need to be solved simultaneously which results from the data-flow between them. The model equations are written to the same MoT file and are solved at the lowest overall time scale instead of solving each scale at its own time-scale. In case numerical problems are encountered and the later performed eigenvalue analysis reveals that the time-scales differ greatly for both scales this decision might need to be revised. Tables 5.17 and 5.18 show the specified variable values for the macro- and micro-scale models, respectively.

Table 5.17 *Variable values for macro-scale, scenario 2 (values taken from Luss & Amundson, 1968)*

dependent	p	0.1 [atm]
	T	600 [°R]
communicated from micro scale	p_p	default
	T_p	default
known	p_e	0.1 [atm]
	T_e	600 [°R]
	T_w	720 [°R]
	H_g	320 [-]
	H_w	1.6 [-]
	H_T	266 [-]

Table 5.18 *Variable values for micro-scale, scenario 2 (values taken from Luss & Amundson, 1968)*

dependent	p_p	0 [atm]
	T_p	600 [°R]
communicated from macro scale	T	default
	p	default
known	A	0.17142 [-]
	C	205.74 [-]
	F	8000 [-]
	H_g	320 [-]
	H_T	266 [-]

The incidence matrix for the multi-scale model is given in Table 5.19.

Table 5.19 *Incidence matrix for scenario 2:*

	$\boldsymbol{p}$	$\boldsymbol{T}$	$\mathbf{k}$	$\boldsymbol{p_p}$	$\boldsymbol{T_p}$	
(5.283)	*			⊛		macro scale
(5.284)		*			⊛	
(5.289)			*		*	micro scale
(5.290)	*		*	*		
(5.291)		*	*	*	*	

The incidence matrix confirms that the macro scale and the micro scale equations are coupled because the values for p_p, T_p, p, and T (see off-diagonal elements) need to be exchanged in each time step.

The construction of scenario 2 is complete. Since also for scenario 2 no parameter estimation is needed the next step is to compare the performance of both scenarios with respect to the modelling objective (Phase IV). Due to the lack of experimental data the two candidate model-scenarios are to be evaluated based on a detailed analysis and comparison of their simulation results (steady state and dynamic behaviour). To create the required simulation results the simulation work-flow is applied (Section 5.3.2). Afterwards, the performance of the scenarios will be compared (Section 5.3.3).

5.3.2 Model application A. Simulation

The different steps of the simulation work-flow are highlighted for the multi-scale scenario 2. For scenario 1 only the main results used for the comparison of the two scenarios are provided in the next section.

Step A.1: Simulation objective

The performance of the fluidized bed reactor is simulated and analysed for the developed two alternative dynamic models (steady state and dynamic behaviour) in order to compare their predictions.

Step A.2: Update variable values

The model variable values provided during the model construction are kept.

Step A.3: Output variables

The desired output variables (plots and values) are the values of the dependent variables, that is, the temperature and partial pressure of the reactant in the fluidized bed and in the catalyst particles. Applying MoT the final values of all model variables are output automatically. For variables that are to be plotted the modeller just needs to select the desired variables from the variable list in order for MoT to generate the corresponding plots during the simulation.

Step A.4: Steady state analysis (differential)

Since the system is dynamic the number and existence of (multiple) steady states are investigated under the relevant conditions. For the resulting steady states the asymptotic stability is checked. This steady state analysis is performed using the steady state model. Afterwards, the convergence to the found steady states is investigated using the dynamic model.

Derive and construct steady state model

The steady state model equations for scenario 2 are given below:

Macro scale:

$$0 = p_e - p + H_g \cdot (p_p - p) \tag{5.295}$$

$$0 = T_e - T + H_w \cdot (T_w - T) + H_t \cdot (T_p - T) \tag{5.296}$$

$$k = 0.0006 \cdot exp\left\{\frac{20.7-15000}{T_p}\right\} \tag{5.297}$$

Micro scale:

$$0 = \frac{H_g}{A} \cdot \left(p - p_p\right) - \frac{H_g \cdot K_k \cdot p_p}{A} \tag{5.298}$$

$$0 = \frac{H_t}{C} \cdot \left(T - T_p\right) + \frac{H_t \cdot F \cdot K_k \cdot p_p}{C} \tag{5.299}$$

Find all steady states/solutions under relevant conditions

All solutions of the steady state model under relevant conditions need to be found. The algebraic solver selected for the solution of the steady state model is the Wegstein solver in MoT. The table below shows the relevant steady states found for scenario 2. They agree with the steady states found by Luss & Amundson (1968).

Table 5.20 *Relevant steady states for multi-scale scenario 2:*

	1st steady state	***2nd steady state***	***3rd steady state***
p	0.09353	0.06704	0.00682
p_p	0.09351	0.06694	0.00653
T	690.445	758.346	912.764
T_p	690.607	759.170	915.094

Investigate asymptotic stability of the steady states and stiffness of system (eigenvalue analysis)

The stability can for example be investigated by calculating the eigenvalues of the Jacobian matrix at the steady state solution. If the real parts of all eigenvalues at the steady state are negative the steady state is stable otherwise it is unstable. For the eigenvalue analysis it is important to consider all ODEs from all scales combined. Table 5.21 summarizes the real parts of the eigenvalues for all three steady states obtained from the eigenvalue report of MoT (automatically created during the solution of the algebraic model applying the Wegstein solver in the previous step) and Figure 5.22 shows a screenshot of the eigenvalue report in MoT for the 1st steady state. For steady states 1 and 3 all eigenvalues have negative real parts. This means that these steady states are asymptotically stable. The second steady state has one eigenvalue with a positive real part and consequently it is unstable. All imaginary parts are 0 which means that there are no oscillations in the system.

Table 5.21 *Real parts of eigenvalues of Jacobian matrix of steady state model for all relevant steady states for multi-scale scenario 2:*

	1st steady state	***2nd steady state***	***3rd steady state***
λ_1	-0.00632	0.00613	-0.06803
λ_2	-0.91232	-1.26657	-13.4706
λ_3	-270.55	-270.55	-269.20
λ_4	-2187.25	-2189.37	-2261.52

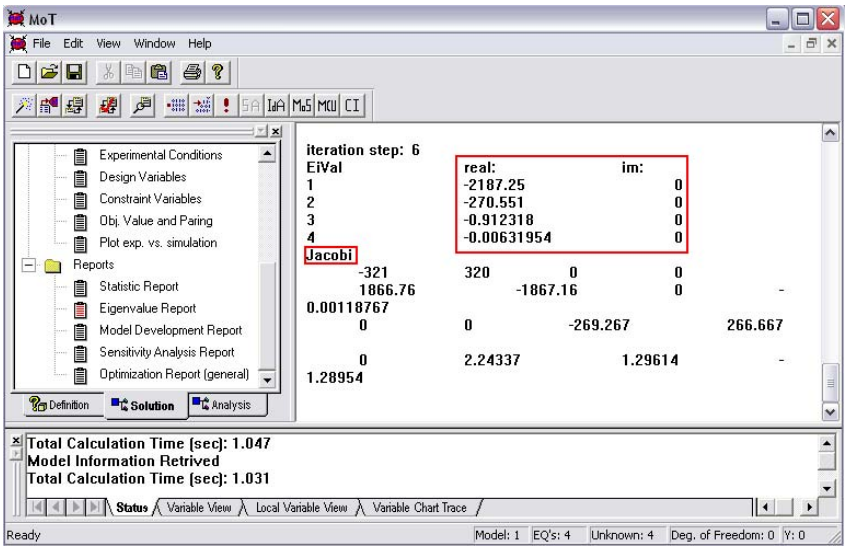


Figure 5.22 *Eigenvalue report for 1st steady state in MoT, scenario 2.*

In the following it will be highlighted why it is important to consider the overall multi-scale model for the eigenvalue analysis instead of performing a separate eigenvalue analysis for each scale. The eigenvalues given in Table 5.21 are obtained when all four model equations of the macro and micro scales are combined as one set of model equations. The Jacobian matrix of this system is a 4x4-matrix and given below for the 1st steady state.

Table 5.22 *Jacobian matrix for 1st relevant steady state for multi-scale scenario 2:*

-321	0	0	320
0	-269.267	266.667	0
0	1.29614	-1.28954	2.24337
1866.76	0	-0.00119	-1867.16

If, however, the micro and macro models are analysed separately there will be one 2x2 Jacobian matrix for each scale. For the 1st steady state these are given in Table 5.23.

Table 5.23 *Separate Jacobian matrices for macro- and micro-scale models for multi-scale scenario 2:*

macro scale:	
-321	0
0	-269.267

micro scale:	
-1.28954	2.24337
-0.00119	-1867.16

The eigenvalues of the micro and macro scale Jacobian matrices are not the same as the eigenvalues of the Jacobian matrix for the overall system because the coupling elements between the scales are missing in the two separated Jacobian matrices. The eigenvalues of the separated micro- and macro-scale models are given in Table 5.24.

Table 5.24 *Real parts of eigenvalues of Jacobian matrices of separate scale models compared to real parts of eigenvalues of overall Jacobian matrix for multi-scale scenario 2 (at 1st steady state)*

	1st steady state	***2nd steady state***	***3rd steady state***	***combined 1st steady state***	
λ_1	-321	-321	-321	-0.00632	***macro***
λ_2	-269.27	-269.27	-269.27	-0.91232	
λ_3	-1867.16	-1869.63	-1949.97	-270.55	***micro***
λ_4	-1.28954	-1.26841	-1.24436	-2187.25	

For the separate scale models it seems that the second steady state is also asymptotically stable. From this it is concluded that for the eigenvalue analysis of a multi-scale scenario always the overall Jacobian matrix of the entire scenario needs to be considered.

The eigenvalue analysis of the entire multi-scale scenario also reveals that the system consists of one very fast mode, one fast mode, one slow mode and one very slow mode. The system is stiff and a dynamic solver that can handle stiffness needs to be applied. Since the smallest eigenvalue is by far smaller than the other eigenvalues the behaviour near the steady state will be mainly determined by this eigenvalue (Luss & Amundson, 1968). The stiffness of the system opens a potential for model reduction by introducing a steady state assumption for the fast modes.

Go back to dynamic model

Output of variables

The dynamic profiles of the dependent variables (partial pressures of reactant and temperatures in gas phase and catalyst particle) are automatically generated by MoT during model solution.

Decide which steady state is to be reached and set initial values of dependent variables accordingly

The dynamic convergence to all detected steady states is investigated. The table below gives the corresponding initial values.

Table 5.25 *Initial conditions for simulation of scenario 2*

	1st steady state	***2nd steady state***	***3rd steady state***
$\boldsymbol{p}$	0.1	0.1	0.1
$\boldsymbol{p_p}$	0	0	0
$\boldsymbol{T}$	600	600	600
$\boldsymbol{T_p}$	690.607	759.170	915.094

These initial conditions are chosen like in (Luss & Amundson, 1968) in order to compare and validate the performance of the constructed models in MoT with the published behaviour.

Step A.5: Run simulation

The BDF solver in MoT has been used for model solution. Table 5.26 shows the final values of the steady states obtained in the dynamic simulations performed (initial values from Table 5.25).

Table 5.26 *Steady state solutions for dependent variables, scenario 2*

	1st steady state	***2nd steady state***	***3rd steady state***
$\boldsymbol{p}$	0.09352	unstable, goes to 1st	0.00669
$\boldsymbol{p_p}$	0.09350	unstable, goes to 1st	0.00639
$\boldsymbol{T}$	690.007	unstable, goes to 1st	913.995
$\boldsymbol{T_p}$	690.165	unstable, goes to 1st	916.342

The dynamic steady state solution is in agreement with the solution of the algebraic steady state model. The convergence to the steady state solutions is shown and discussed in more detail:

1st steady state:

For the temperatures of the fluidized bed and the catalyst particles the 1st steady state is reached for a dimensionless time τ of 0.02. The partial pressure of the reactant converges to the 1st steady state by a factor of 10 faster (τ=0.002). Figure 5.23 shows the partial pressures and temperatures in the bulk and the catalyst particles with respect to time.

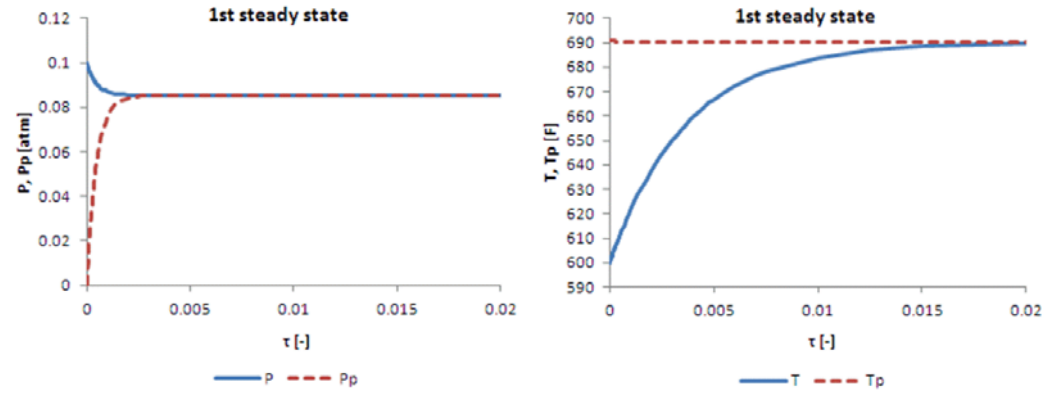


Figure 5.23 *Convergence of reactant partial pressures (left) and temperatures (right) in fluidized bed and catalyst particles to 1st steady state for scenario 2.*

These results confirm the outcome of the eigenvalue analysis of the steady state model. The system consists of two fast modes, which apparently are the partial pressures, one slow mode and one very slow mode. The temperatures of catalyst and bulk correspond to the slow modes whereas the dynamic response of the pellet temperature is especially slow. The main reason for the slow dynamic response of the catalyst temperature is its high heat capacity (Luss & Amundson 1968). The partial pressure of the reactants in the catalyst particles rises very fast until it equals the reactant partial pressure in the bulk. It can be concluded that the reaction is the rate-limiting step and not the mass transfer. Consequently, it might be possible to reduce the model for scenario 2 by neglecting the mass transfer resistance between bulk gas and catalyst pellet. The same seems to hold true for the heat transfer resistance between gas bulk and pellet because the bulk gas temperature reaches the value of the catalyst temperature relatively fast.

2nd steady state:

The second steady state is also reached faster for the partial pressures than for the temperatures. The particle and gas phase partial pressures reach the steady state after a τ of approximately 0.002 whereas the gas phase temperature reaches steady state for a τ of 0.02. The dynamic behaviour is shown in Figure 5.24.

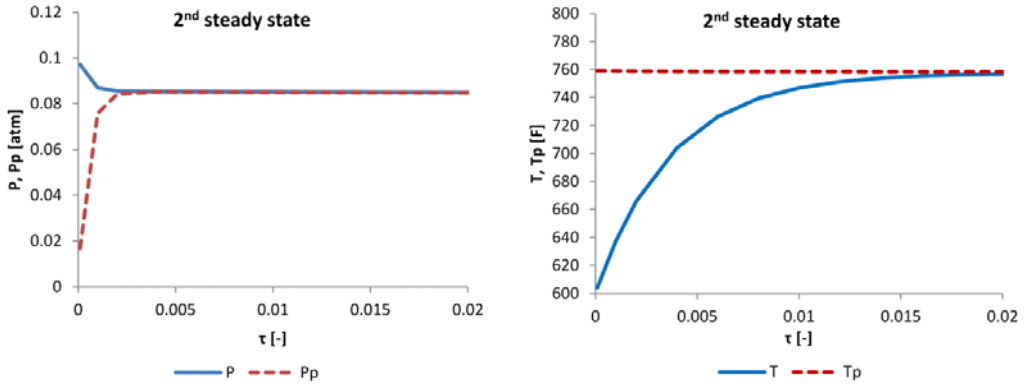


Figure 5.24 *Convergence of reactant partial pressures (left) and temperatures (right) in fluidized bed and catalyst particles to 2nd steady state for scenario 2.*

Furthermore, the dynamic simulations confirm the instability of the 2nd steady state identified during the eigenvalue analysis of the steady state model. After a τ of around 100 a transition between the unstable second steady state and the stable 1st steady state occurs which is reached at a τ of about 1000 (see Figure 5.25).

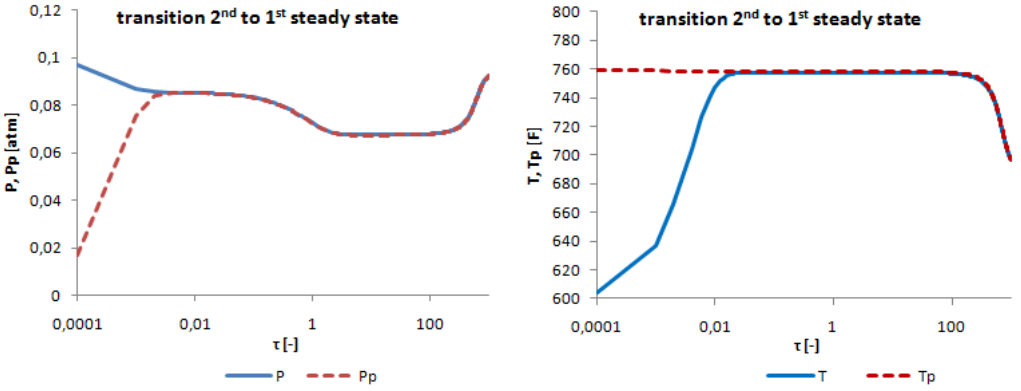


Figure 5.25 *Convergence of reactant partial pressures (left) and temperatures (right) in fluidized bed and catalyst particles from 2nd to 1st steady state for scenario 2.*

3rd steady state:

The 3rd steady state is reached for a τ of about 0.2. Figure 5.26 shows the plot of the temperatures and pressures with respect to time.

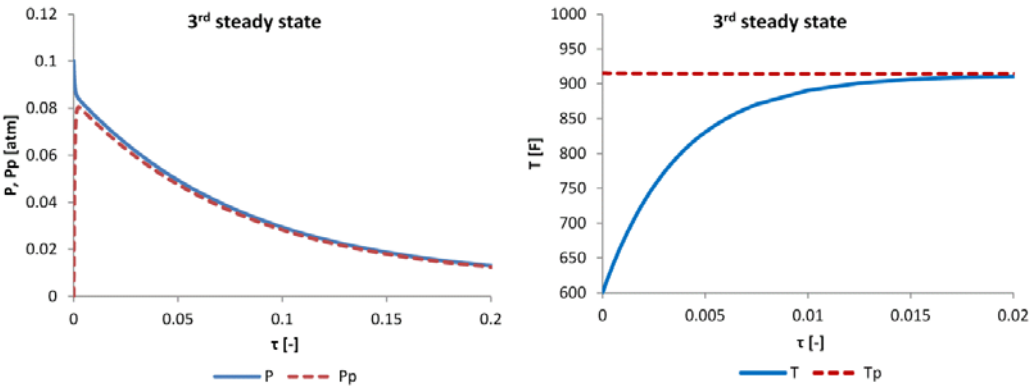


Figure 5.26 *Convergence of reactant partial pressures (left) and temperatures (right) in fluidized bed and catalyst particles to 3rd steady state for scenario 2.*

Here, the partial pressures reach steady state later than the gas phase temperature.

For all 3 steady states the resulting behaviour is in good agreement with the published results (Luss & Amundson, 1968) for this multi-scale scenario.

The simulation and evaluation for scenario 1 is performed in the same way like for scenario 2.

5.3.3 Model evaluation/ validation (Phase IV)

The performance of scenarios 1 and 2 is compared through the simulation results. Scenario 2 has three different steady states that have been given in Table 5.20. The first and the third steady state in Table 5.20 have resulted to be asymptotically stable in the eigenvalue analysis. The steady state behaviour of scenario 1 differs significantly from scenario 2. Only one stable steady state exists at p=0.000666 atm, T=927.91 F. Since already the steady state behaviour of the simpler scenario 1 differs from that of scenario 2 it can be concluded that the assumptions made for scenario 1 are not valid. The existence of the catalyst particles cannot be neglected. Consequently it is not necessary to compare the dynamic behaviour.

The eigenvalue analysis for scenario 2, as well as the plots of the dynamic behaviour of the system (Figures 5.24-5.26) suggest, however, that the model for scenario 2 might be reduced by neglecting the mass and heat transfer resistances between the catalyst particles and the gas bulk because heat and mass transfer are first compared to the reaction and the temperature change of the catalyst pellets. These consideration lead to the development of a third scenario (scenario 3) which neglects the heat and mass transfer resistances between catalyst and gas bulk. The performance of scenario 3 (steady state and dynamic behaviour) is compared to the more complex scenario 2. In contrast to scenario 1 the existence of the particles is still considered in scenario 3.

5.3.4 Model reduction based on eigenvalue analysis and evaluation of dynamic simulations

Scenario 3 can be derived from scenario 2 by setting p_p=p and T_p=T (Luss & Amundson, 1968). The resulting model equations for scenario 3 are given by:

$$\mathrm{k} = 0.0006 \cdot exp\left\{\frac{20.7-15000}{T}\right\} \tag{5.300}$$

$$\frac{dp}{dt} = (p_e - p - H_g \cdot K_k \cdot p)/(1 + A) \tag{5.301}$$

$$\frac{dT}{dt} = (T_e - T + H_w(T_w - T) + H_T \cdot F \cdot K_k \cdot p)/(1 + C) \tag{5.302}$$

Here, a single-scale scenario is being considered. The incidence matrix for scenario 3 is given in Table 5.27. It is divided in the AE and ODE parts. It can be seen that the algebraic equation is coupled to the ODE-part and cannot be solved separately. The ODE-part is of lower tridiagonal form.

Table 5.27 *Incidence matrix for scenario 3*

	K_k	p	T	
(5.300)	*		⊛	***AE-part***
(5.301)	*	*		***ODE-part***
(5.302)	*		*	

In order to compare the performance of scenario 3 with scenario 2, first the steady state behaviour is compared. Scenario 3 has the same steady states as scenario 2. Table 5.28 summarizes the real parts of the eigenvalues for both scenarios at the three relevant steady states.

Table 5.28 *Real parts of eigenvalues at steady state for scenarios 3 and 2*

	1st steady state	***2nd steady state***	***3rd steady state***
		Scenario 3	
λ_1	-0.00651264	0.00620247	-0.0088858
λ_2	-0.911608	-1.28696	-12.3175
		Scenario 2	
λ_1	-0.00632	0.00613	-0.06803
λ_2	-0.91232	-1.26657	-13.4706
λ_3	-270.55	-270.55	-269.20
λ_4	-2187.25	-2189.37	-2261.52

Table 5.28 reveals that after the undertaken simplification from scenario 2 to scenario 3 only the slow modes are left. The fast modes were represented by the mass transfer that is now assumed to be infinitely fast and thus is no longer considered in the dynamic behaviour. The eigenvalues further reveal that the asymptotic stability of the steady states in scenario 3 remains unchanged in comparison to scenario 2. That means that the 1st and the 3rd steady states are asymptotically stable whereas the 2nd steady state is unstable.

In a second step, the dynamic behaviour of the scenarios is compared. The convergence for both scenarios to all three identified steady states is shown in Figure 5.27.

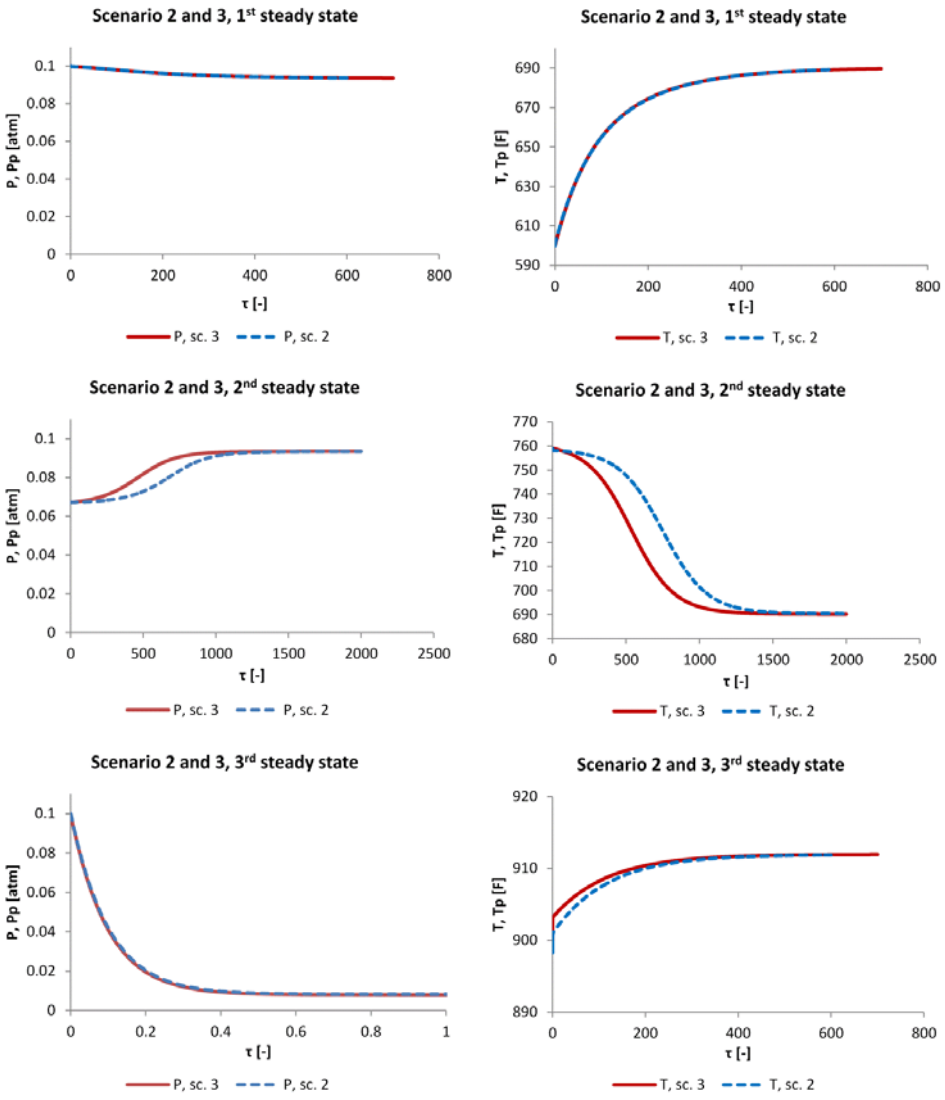


Figure 5.27 *Convergence of reactant partial pressures (left) and temperatures (right) in fluidized bed and catalyst particles to 3rd steady state for scenario 2.*

The figure shows that the dynamic profiles of both scenarios are very similar. However, especially for the convergence from the second to the third steady state there are some deviations between scenario 2 and 3. The path for reaching the 3rd steady state is different even though the same steady state is reached. Since for this simple example the computational cost for the more complex scenario 2 is not noticeably higher, it is recommended to use scenario 2 for dynamic simulations of the fluidized bed reactor.

5.3.5 Conclusions for case study

The case study highlights the application of the top-down strategy together with the multi-scale model construction work-flow to derive a multi-scale scenario from a simple single-scale starting scenario. Furthermore, the simulation work-flow for dynamic models is verified and it is illustrated how the eigenvalue analysis can be applied to evaluate the asymptotic stability of steady states. In addition, the importance of the eigenvalue analysis in identifying a potential for model reductions is demonstrated. Three different model-scenarios for the fluidized bed reactor have been compared based on a detailed simulation analysis. The simplest scenario is the starting scenario which considers only the fluidized bed scale and neglects the existence of the catalyst particles. The second scenario has been derived from the first by increasing the degree of detail and including an additional scale for the catalyst particles. The third scenario has resulted from a reduction of the second scenario which was indicated by the eigenvalue analysis and the analysis of the dynamic profiles of scenario 2. In scenario 3 the mass and heat transfer resistance between gas bulk and catalyst pellets is neglected which converts scenario 2 to a single-scale scenario. In contrast to scenario 1 the resulting scenario 3 does not neglect the existence of the catalyst particles. It turns out that scenario 1 is not able to capture the steady state behaviour of the system. Scenario 2 has been used as a reference because it is the most complete. Scenario 3 captures both the steady state and the dynamic behaviour. However, in some cases deviations occur for the dynamic behaviour compared to the reference scenario 2.

5.4 Fragrance industry: Fragrance aerosol system

The development of the computer-aided modelling framework (Chapter 3) has revealed that a modelling tool needs to supply multi-scale templates which provide specific support for a class of similar problems. In this context, a template is defined as a computer-aided work-flow for a specific problem which is based on the generic computer-aided work-flows presented in Chapters 2 and 3 but in addition provides domain-specific support for each step of the generic work-flows.

To address this need a modelling template for the prediction of the fate of a fragrance aerosol under the effects of sedimentation, dispersion, evaporation, agglomeration and breakage is proposed here. The template supports the systematic derivation of appropriate models, with respect to a specific modelling goal and available experimental data. The fragrance spraying template is a template within the generic work-flows for the 'Modelling objective and system documentation'-phase and the 'Multi-scale model construction'-phase. It serves only as one example on how domain-specific templates can be integrated and are compatible with the generic work-flows proposed in Chapter 2 and the resulting computer-aided modelling framework (Chapter 3).

The fragrance aerosol system has been chosen because it is challenging and of high interest for the related industries. The developed models are to be applied to predict, evaluate and optimize the fragrance product (e.g. air freshener, fine fragrance) qualities and performance. On the one hand, models that predict the initial size distribution of the created droplets during a spraying process based on different system properties like spray-can pressure, propellants, active ingredients and the type of the spraying device are of interest. On the other hand, there is a need for models that are able to describe the fate of the created droplets - for example the settling and distribution in space compared to evaporation or the size and composition of the evaporating stream from the droplets with respect to position and time. For the developed template the focus is on the 2nd part; the modelling of the fate of the droplets after they have been generated. A population balance describing the droplet size distribution is applied. Different phenomena like dispersion, sedimentation, convection, evaporation, agglomeration and breakage influence the fragrance droplet size distribution and the relevant terms need to be included in the population balance model. In the literature a number of different models for the evolution of droplet size distributions are available for different systems. Simon et al. (2003), for example, have proposed a droplet population balance model including phenomena like axial dispersion, droplet rising, breakage and coalescence for the modelling of the hydrodynamic behaviour of solvent extraction columns. Koch has proposed a model for aerosols created during a spraying process that contains a simple evaporation model together with sedimentation and

dispersion. What is still missing however is a systematic strategy for fragrance aerosol modelling that is generic and provides:

- a systematic strategy to derive specific fragrance spraying models from a generic model;
- a tool for retrieval of needed models from a model library and strategies for linking them;
- methods for model analysis, identification, discrimination and solution;
- links to databases for fragrance compounds;
- a strategy to derive the needed property models for fragrance compounds;
- integration of the above within a computer-aided modelling framework;

The developed fragrance aerosol template aims to tackle these problems in combination with the generic computer-aided modelling framework. The fragrance aerosol template and its integration in the generic modelling framework is described in detail in Section 5.4.1 and is further highlighted by a case study in Section 5.4.2. The case study is related to the derivation of a fragrance aerosol model that is able to reflect measured dynamic droplet size distribution profiles for limonene. These two sections have been written based on a submitted manuscript (Heitzig et al., 2012). In Section 5.4.3 a second case study is presented briefly. In contrast to the first case study the second case study is not related to the development but to the application of a developed fragrance aerosol model. The second case study has also been published before (Heitzig et al., 2011b). Finally, in Section 5.4.4 the main conclusions are summarized.

5.4.1 Fragrance aerosol template

Figure 5.28 shows the fragrance aerosol template superimposed on the different work-flow steps for Phases I and II.

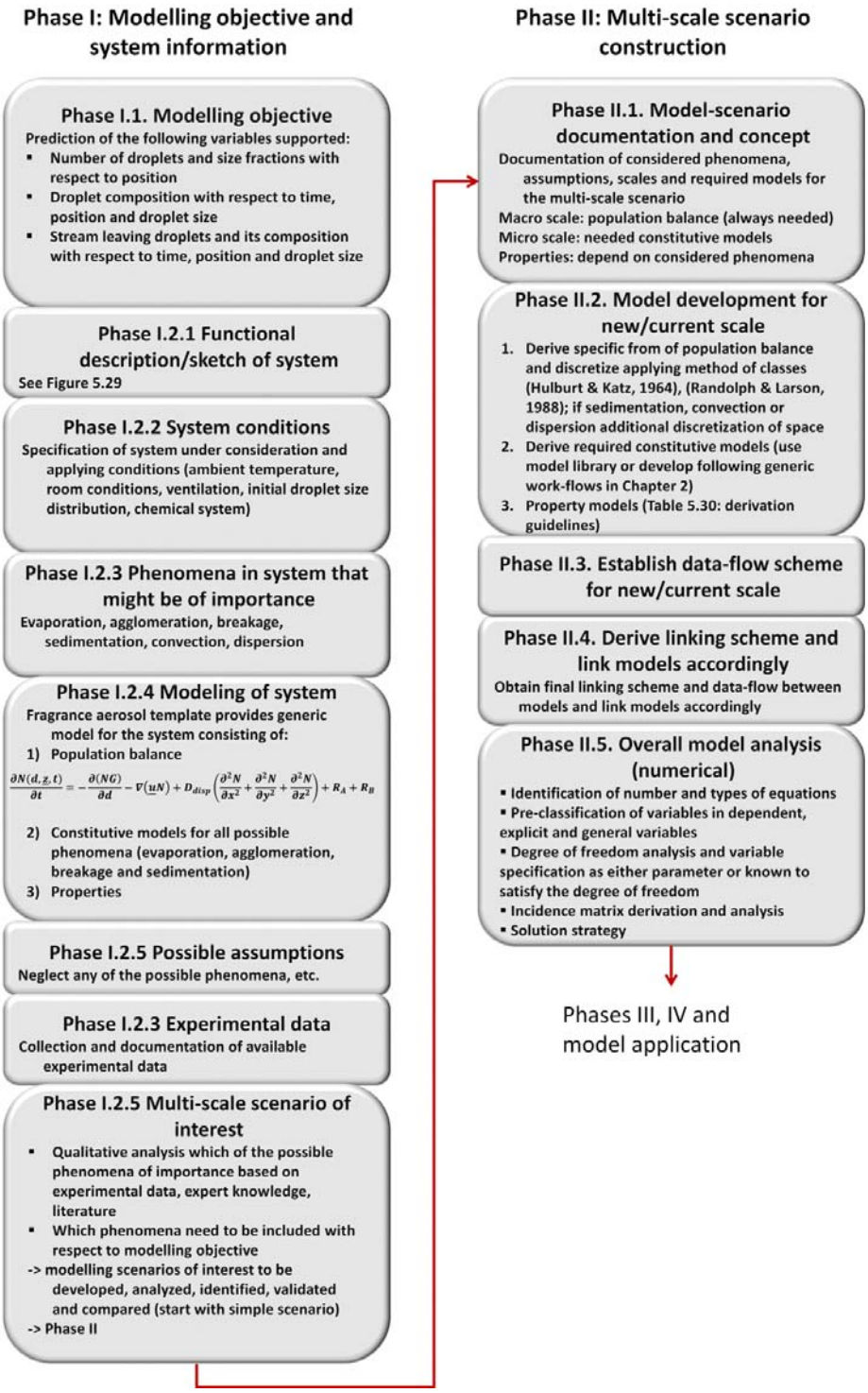


Figure 5.28 *Presentation of fragrance aerosol template for each step of generic work-flows for 'modelling objective and system information' (Phase I) and 'Multi-scale model construction' (Phase II).*

Step I.1: Modelling objective (Phase I)

The generic model allows the derivation of different models for the fragrance system. These models differ in the behaviour they are able to predict, considered phenomena and in degree of detail of the models for the different phenomena.

The prediction of the following variables is supported:

- Number of droplets and size fractions with respect to time and position,
- Droplet compositions with respect to time, position and discrete droplet size,
- Stream leaving droplets and its composition with respect to time, position and discrete droplet size.

Step I.2: System information and documentation (Phase I)

The general information available on the fragrance aerosol system is summarized. For different applications this information might need extension.

Step I.2.1: Functional description/ sketch of the system to be modelled (process, unit operation, product)

Figure 5.29 gives a general sketch of the fragrance spraying system and the main phenomena that can be of importance and are therefore supported.

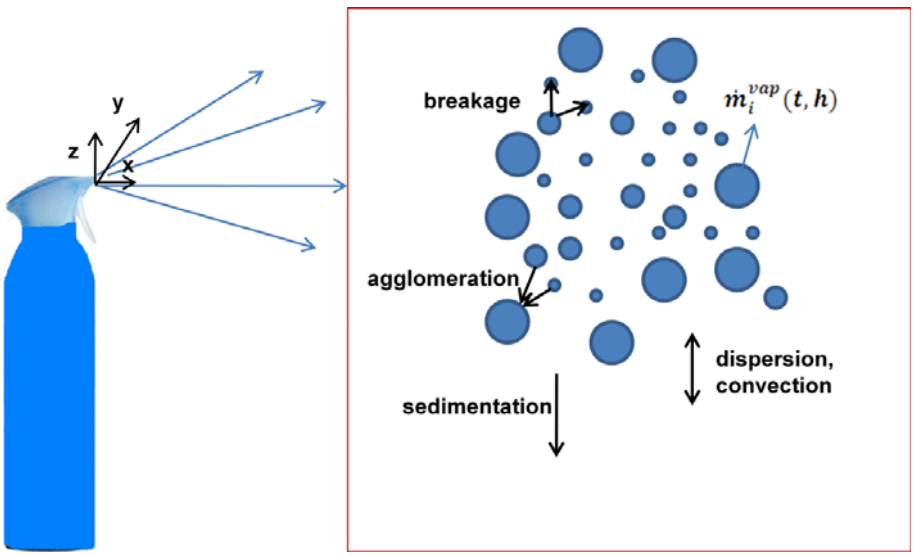


Figure 5.29 *Sketch of the fragrance aerosol system.*

Step I.2.2: System conditions

It is important to specify the system under consideration and the conditions that apply. What definitely needs to be clarified are the ambient temperature, the room conditions in general (e.g. strong or no ventilation), the initial droplet size distribution and the chemical system.

Step I.2.3: Phenomena in the system that might be of importance

The phenomena that might be of importance for the fragrance aerosol system are depicted in Figure 5.29. These are: evaporation, agglomeration, breakage, sedimentation, dispersion and convection.

Step I.2.4: Modelling of system/problem

In this step all available information on how the system under consideration can be modelled is collected and documented. The fragrance spraying template provides a modelling framework, where from a generic model, problem specific models are generated. This is achieved by selecting problem specific options within the generic model in terms of balance equations and the corresponding constitutive models and equations. Population balance models as well as alternative constitutive models for the different phenomena are available in the model library.

A) Balance equations

Equation 5.303 shows the generic form of the population balance for the droplet size distribution.

$$\frac{\partial N(d,\underline{z},t)}{\partial t} = -\frac{\partial (NG)}{\partial d} + \nabla\left(\underline{u}N\right) + D_{disp}\left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} + \frac{\partial^2 N}{\partial z^2}\right) + R_A + R_B \tag{5.303}$$

Here, N is the concentration of the droplets with respect to droplet diameter d, droplet position $\underline{z}$ and time t. G is the growth rate of the droplets, $\underline{u}$ is the droplet velocity vector, D_{disp} is the dispersion coefficient. R_A and R_B are the agglomeration and breakage rates, respectively. In case the droplet transport due to convection is neglected the general population balance transforms to:

$$\frac{\partial N(d,\underline{z},t)}{\partial t} = -\frac{\partial (NG)}{\partial d} + v_{sed}(d)\frac{\partial N}{\partial z} + D_{disp}\left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2} + \frac{\partial^2 N}{\partial z^2}\right) + R_A + R_B \tag{5.304}$$

The population balance needs to be converted to a solvable form applying either the method of moments or the method of classes. The method of classes (Hulburt & Katz, 1964; Randolph & Larson, 1988) is used because it is applicable for time-dependent growth rates G. This method is based on discretizing the droplet diameter in size fractions.

An additional energy balance is not considered because it is assumed that the ambient temperature surrounding the droplets does not change significantly due to droplet evaporation. Furthermore, it is assumed that the fragrance concentration in the surrounding gas far away from the droplet is not affected by droplet evaporation.

B) Constitutive models and equations

Depending on the problem-specific form derived from the generic population balance different constitutive models and equations need to be considered.

B1) Model for growth rate

The droplets decrease in size due to evaporation. Droplet evaporation models have been developed by a number of authors. Different alternatives based on models proposed by Renninger et al. (1981), Kukkonen et al. (1989), Ranz & Marshall (1952) have been introduced to MoT and extended according to the desired degree of detail. The different alternative evaporation models have been validated with experimental data from Ranz & Marshall (1952) for the evaporation of pure water droplets without convection and are available in the library. Table 5.29 gives an overview of the assumptions all these models incorporate and of the different degrees of details that can be considered.

Table 5.29 *Possible degree of detail for evaporation model and assumptions*

Possible degree of detail	***Assumptions***
▪ Dynamic or steady state energy balance ▪ Vapour pressure at equilibrium at droplet surface ▪ Kelvin effect (vapour pressure around droplet compared to over plane surface) ▪ Ideal or non-ideal mixtures ▪ Diffusion from droplet in surrounding gas ▪ Heat conduction from droplet in surrounding gas ▪ Different assumptions about concentration and temperature profiles in surrounding gas in radial direction (e.g. linear) ▪ Convection (effect on mass and heat transport in surrounding gas) ▪ Heat and mass transfer resistance at droplet surface ▪ Radiation ▪ Gas kinetic transfer	▪ Droplets are always spherical ▪ Vapour liquid equilibrium (VLE) is established at droplet surface ▪ Vapour behaves like an ideal gas ▪ Effect of other droplets and droplet evaporation negligible ▪ Diffusivity through air of each vapour compound is unaffected by the presence of the other vapour compounds ▪ Neglect Stefan flow ▪ Droplets are ideal mixed

The modeller is not obliged to make the assumptions listed in Table 5.29. However, in case some of the assumptions should not be made the corresponding model is not available from the library and an existing library model needs to be extended accordingly.

For the evaporation models different properties are needed. The properties and/or property models can be derived from the model library or from the connected thermodynamic library of ICAS. In case a property is not available the gap needs to be filled using values or data given in

literature. In case the required information is not available in open literature, experiments need to be performed. For pure compound properties the gap can also be filled through property prediction models. The software ICAS contains a tool for group and atom contribution based property prediction called ProPred (Gani et al., 1997). Table 5.30 lists the required properties and their recommended sources.

Table 5.30 *Required properties and their models/sources*

Properties	***Examples***	***Source***
Pure compound properties		1. thermodynamic database
constant	T_c, MW	2. literature
T-dependent	$cp_{vap,i}(T)$, $cp_{liq,i}(T)$, $L_i(T)$, $P_i^{s,plane}(T)$, $\sigma_i(T)$, $\rho_i(T)$, $D_{air}(T)$, $K_{air}(T)$	3. a) experiments b) property prediction
Mixture properties	Activity coefficients (T, C_a)	1. thermodynamic model library (models and coefficients) 2. literature 3. UNIFAC

B2) Agglomeration model

The proposed agglomeration model considers agglomeration of only two droplets at a time. A model for the agglomeration rate $R_{A,i}$ for a discrete droplet size i has been proposed by Costa et al. (2005):

$$R_{A,i} = \sum_{l=1}^{Ndis(Ndis+1)/2} \upsilon_{l,i} \cdot r_A(l) \tag{5.305}$$

Here, $Ndis$ is the number of discrete diameters in the droplet size distribution, l is summed over all possible agglomerations between two discrete droplet sizes n and m of the droplet size distribution and $r_A(l)$ is the agglomeration rate of the agglomeration with index l. The stoichiometric factor of the agglomeration l with respect to the discrete droplets size i is given by $\upsilon_{l,i}$:

$$\upsilon_{l,i} = \left(\frac{S_m^3+S_n^3}{S_q^3}\right) \cdot \delta_{i,q} - \left(\delta_{i,m} + \delta_{i,n}\right) \tag{5.306}$$

S_m and S_n are the mean diameters of the two agglomerating size fractions m and n of the agglomeration with index l whereas S_q is the mean diameter of the created size fraction by agglomeration l. They are calculated as follows:

$$S_i = \frac{d_{i-1}+d_i}{2} \tag{5.307}$$

The factors $\delta_{i,q}$, $\delta_{i,m}$ and $\delta_{i,n}$ are 0 except for the case when i is q, m or n, respectively. With this generic stoichiometric factor it is assured that $R_{A,i}$ combines the rates of all agglomerations in which droplets of size i are formed or destroyed with the corresponding sign.

An expression for the rate of agglomeration with index l where droplets n and m agglomerate to form the droplet q has been adapted from Coulaloglou & Tavlarides (1977). These authors have derived the rate expression for the agglomeration of droplets in mixing reactors. The rate expression depends on the concentration of the agglomerating droplets, the agglomeration frequency and the agglomeration efficiency:

$$r_A(l,q) = r_A(n,m,q) = h(n,m) \cdot \lambda(n,m) \cdot N_n \cdot N_m \quad (5.308)$$

N_n and N_m are the concentrations of the agglomerating droplets of size fractions n and m. The agglomeration frequency $h(n,m)$ and the agglomeration efficiency $\lambda(n,m)$ are given by:

$$h(n,m) = C1(S_n^2 + S_m^2) \cdot \left(S_n^{2/3} + S_m^{2/3}\right)^{1/2} \quad (5.309)$$

$$\lambda(n,m) = exp\left\{-\frac{C2}{\sigma_i^2}\left(\frac{S_n S_m}{S_n + S_m}\right)^2\right\} \quad (5.310)$$

Here, σ_i is the surface tension of the droplets with discrete diameter d_i. $C1$ and $C2$ are adjustable model parameters. The expressions for the frequency and efficiency given in Equations 5.309 and 5.310 have been modified from the ones proposed by Coulaloglou & Tavlarides (1977) because the original equations depend on a number of properties which are related to the conditions surrounding the droplets in the mixed reactor, e.g. the agitator speed, which do not play a role for the fragrance aerosol system. These properties have been combined with the adjustable parameters C1 and C2 because in the aerosol system the impact of the surrounding environment on the droplets is very different from the reactor and depends on the conditions in the room where the droplets have been sprayed into. For the same room or similar room conditions this impact is the same for different fragrance systems and therefore it makes sense to merge the impact or the surrounding room conditions with the adjustable parameter of the agglomeration model.

B3) Breakage model

The form of the breakage model is similar to the agglomeration model. It is assumed that a droplet always breaks in not more than two smaller droplets. The breakage rate $R_{B,i}$ for a discrete droplet size i is given by:

$$R_{B,i} = \sum_{l=1}^{Ndis(Ndis+1)/2} \upsilon_{l,i} \cdot r_B(i) \quad (5.311)$$

An expression for the breakage rate $r_B(i)$ for the droplets of discrete mean diameter $S[i]$ ·has been adapted from (Cauwenberg, 1995; Bahmanyar & Slater, 1991) for agitated liquid-liquid dispersions. The rate expression has been adjusted for the aerosol system in a similar way like the agglomeration rates.

$$r_B(i) = -\frac{C3 \cdot S[i] \cdot N[i]}{\sigma_i} \quad (5.312)$$

Here, $C3$ is an adjustable parameter. The stoichiometric factor $\upsilon_{l,i}$ is given by:

$$v_{li} = \delta_{i,q} - \left(\frac{S_q^3}{S_m^3+S_n^3}\right)\left(\delta_{i,m} + \delta_{i,n}\right) fr_{q\to n+m} \tag{5.313}$$

The factor $fr_{q\to n+m}$ quantifies the portion of all breaking droplets from size fraction q that form droplets of size fractions n and m. It is assumed that the likelihood of the diameter of the created droplets follows a beta-distribution which for the case of breakage into two droplets is given by:

$$\beta(S_0, S_i) = 6 \cdot \frac{{S_i}^5}{S_0^6} \tag{5.314}$$

From Equation 5.314 the likelihood that two droplets of size fractions n and m are formed from breakage of a droplet of size fraction q is given by the evaluation of the double integral given below:

$$fr^*_{q\to n+m} = \int_{d1=d_{m,min}}^{d_{m,max}} \int_{d2=d_{n,min}}^{d_{n,max}} 36 \cdot \frac{d1^5 \cdot d2^5}{S_0^{12}} dd2\ dd1 \tag{5.315}$$

Equation 5.315 does not consider that it is geometrically not possible to form all combinations of m and n by breakage of q into two droplets. For that reason, in order to obtain $fr_{q\to n+m}$ a normation is required where $fr^*_{q\to n+m}$ is divided by the sum of $fr^*_{q\to i+b}$ for all possible pairs of size fractions that can be formed by breakage of droplets from size fraction q.

B4) Sedimentation velocity equation

The sedimentation velocity is calculated depending on the Reynolds number by either Stokes law, Newton's law or a transition law between Stokes regime (Stokes' law is valid) and a Newton regime (Newton's law is valid).

Step I.2.5: Assumptions

Assumptions supported by the modelling template are:

- Neglect change in temperature of air/gas surrounding the droplets;
- Neglect change of concentration of droplet compounds in surrounding far away from droplets;
- Neglect any of the phenomena shown in Figure 5.29 (such as dispersion, evaporation, breakage, etc.)

The specific assumptions for the different phenomena models are not provided here but together with the corresponding models.

Step I.2.6: Preliminary system data

The available experimental data is collected and documented.

Step I.2.7: Select model scenarios of interest

In this step it is analysed which of the possible phenomena (evaporation, agglomeration, breakage, sedimentation, dispersion, convection) might be of importance. This can be done qualitatively based on the available experimental data and based on knowledge from literature or experts. Furthermore, it needs to be identified which phenomena should be included in the model with respect to the modelling goal. If for example one of the modelling goals is to predict the composition and size of the stream evaporating from the droplets, obviously, evaporation needs to be included in the model. Based on this analysis, different modelling scenarios of interest are derived, which in the later steps are systematically constructed, analysed, identified, validated and compared with each other. In general, it is good practice to start with a simple scenario and then, if necessary, gradually increase the degree of detail.

With this last step the system information and documentation part is completed and the modeller can start to develop the models for the selected scenarios of interest. For this purpose the multi-scale model construction work-flow of the generic computer-aided modelling framework is employed. The fragrance spraying template also provides domain-specific support in the different steps of the model construction, which is highlighted below.

Step II.1: Model-scenario documentation and concept (Phase II)

The considered phenomena, assumptions, scales and required models for the multi-scale scenario are documented.
For the spraying process the problem-specific form of the population balance is always needed. The population balance model forms the macro-scale because it considers the entire droplet size distribution.
For the phenomena considered, a corresponding constitutive model is required, which might add new scales to the scenario. The droplet evaporation phenomenon, for example, focuses on single droplets instead of the overall droplet size distribution.
Also, property models may be necessary.

Step II.2: Model development for new/current scale (Phase II)

In this step the required models for all selected scales are developed. Firstly, the problem-specific population balance needs to be derived from the generic population balance and discretized applying the method of classes. In case sedimentation is considered the population balance in addition needs to be discretized with respect to height and if convection or dispersion

are considered, the population balance needs to be discretized in all spatial directions in which dispersion or convection are considered.
Secondly, all required constitutive models for the considered phenomena need to be derived. The modeller has three options:

1. Derive model from fragrance spraying template model library;
2. Derive model from fragrance spraying template model library but further modify model;
3. Develop a new model by going back to Phase I, Phase II and, if needed, Phase III and IV of the generic computer-aided modelling framework.

It is advisable to start with relatively simple models for the different phenomena and to return to this step and refine the models in case their performance with respect to the modelling goal is not satisfactory. Sensitivity analysis is a useful tool to identify the importance of the different components of the model.
Finally, different properties are required depending on the different phenomena considered and their corresponding models. The property models should be connected from the thermodynamic library of ICAS or from the MoT model library, which contains property models for different fragrance compounds. A list of properties that may be needed and guidelines for their derivation is given in Table 5.30.

Step II.3: Establish data-flow scheme for new/current scale (Phase II)

For each developed constitutive/scale model the data-flow between the different models within the multi-scale scenario needs to be updated.

Step II.4: Derive linking scheme and link models accordingly (Phase II)

In this step the final linking scheme and the data-flow between the different models is derived and the models are linked accordingly.

Step II.5: Overall model analysis (numerical) (Phase II)

A numerical model analysis is performed and the solution strategy of the model is derived. A detailed explanation of the numerical model analysis is given in Chapter 2 (Phase II.A+B: Single- and multi-scale model construction work-flows).

MoT provides strong support for fragrance aerosol modelling. The model library contains models (including documentation) for the different phenomena that might be of importance (see Figure 5.29). Evaporation models of differing degree of detail are available (according to Table 5.29). In addition, alternative overall model scenarios of differing degree of detail

including population balances, corresponding constitutive models and property models have been developed. Models for the required properties (vapour pressure, surface tension, activity coefficients, etc.) have been derived from a literature review and property prediction and implemented in MoT for a number of fragrance and propellant compounds because these compounds are not available in the CAPEC-database in ICAS.

5.4.2 Fragrance aerosol case study 1

The systematic derivation of a fragrance aerosol model that is able to reflect observed dynamic droplet size distribution profiles for limonene in MoT according to the proposed fragrance aerosol template (see Figure 5.28) is highlighted in this section.

Step I.1: Modelling objective (Phase I)

The modelling objective is to:

- Systematically derive a problem-specific model from the generic fragrance spraying model that reflects the real system behaviour (represented by measured data from experiments). In order to achieve this it needs to be identified which phenomena have relevance for the investigated system under the conditions of the experiments.
- Use the available experimental data to identify and validate the problem-specific model.

Step I.2: System information and documentation (Phase I)

Step I.2.1: Functional description/ sketch of the system to be modelled (process, unit operation, product)

A general sketch of the system to be modelled with the phenomena that are being considered is shown in Figure 5.29.

Step I.2.2: System conditions

The ambient temperature is 297.15 K. There is no significant ventilation in the room. The initial droplets size distribution is known from the available experimental data and the simulations need to be conducted for the experimental set-up.

Step I.2.3: Phenomena in the system that might be of importance

The potentially relevant phenomena are summarized in Figure 5.29.

Step I.2.4: Modelling of system/problem

In this step the modeller collects information on how the system can be modelled and how the system has been modelled before (e.g. in literature). As shown in the previous section the fragrance spraying template provides a generic population balance based model of the system with associated constitutive models for the possible phenomena in the system. Models for the required properties can be taken from the model library and from connected thermodynamic libraries and property prediction tool.

Step I.2.5: Possible assumptions

Possible assumptions are given by the fragrance aerosol template (see Figure 5.28).

Step I.2.6: Preliminary system data

Experimental data is available for the fate of a limonene aerosol after the spraying process. Light scattering experiments have been performed by Firmenich (for whom this model is being developed) to record the transmission and droplet size distribution of the limonene aerosol with respect to time applying the laser diffraction particle size technique (Spraytec) from Malvern. Figure 5.30 shows the measurement of data.

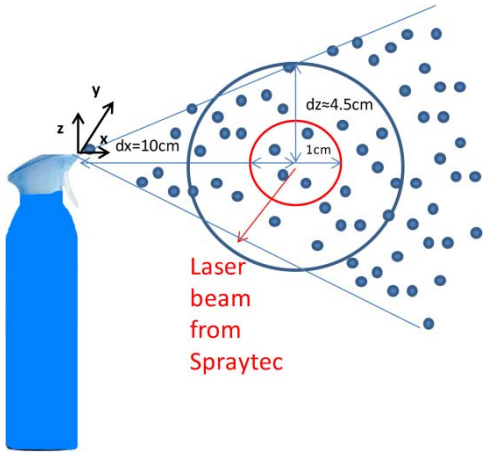


Figure 5.30 *Measurement of data including distance of measurement zone (dx=10 cm) from spray nozzle, laser beam diameter (diameter=1 cm) and diameter of resulting spraying circle in distance (dx) of measurement zone (diameter=9 cm).*

The goal of the experiment was to detect the fate of the created limonene aerosol, that is, the changes of the droplet size distribution with respect to time after the spraying process.
For the experiment, 6 g of the limonene was filled into a spray-can. Nitrogen (0.5 g) was used as a propellant and was added to the can (volume 5.915×10^{-5} m^3) until a pressure of 7.03 atm was reached. Figure 5.30 shows the spraying operation and the droplet distribution that is measured

including the resulting spraying cone and the measurement zone. The laser beam for the light scattering measurements was positioned perpendicular to the spraying direction (x-direction) with a distance of dx=10 cm from the nozzle tip of the spraying can. The diameter of the laser beam was 1 cm. The measurements were started when the spray button was pressed. After approximately 1 second, the button was released. The actual data was recorded after the impact of the spraying process had disappeared.

In addition to the light scattering experiments, it was also determined how wide the droplets are distributed in the plane of the measurement zone (distance dx= 10 cm) by placing a paper in the same distance and measuring the diameter of the spray circle on this paper. The centre of this spray circle was in the centre of the measurement zone and its radius was 4.5 cm (see Figure 5.30). During the spraying experiments there was no significant ventilation within the room. The detected droplet size distribution with 31 discrete diameter fractions having diameters between 1.166 µm and 135.936 µm is given in Appendix A5.

Step I.2.7: Select model scenarios of interest

In this step it is identified which phenomena occur for sure, which phenomena are likely to have a strong impact on the development of the droplet size distribution with respect to time and which phenomena are of special interest with respect to the modelling objective. Furthermore, it is identified which phenomena play no role. At this stage, the analysis is rather qualitative and might need to be revised in case none of the simulated scenarios can be validated by the available experimental data.

Figure 5.29 gives on overview of the different phenomena that might play a role in the described system behaviour. The evaporation is known to occur for sure (from available expert knowledge and thermodynamics) although at this point it is not known if this phenomenon has significance for the development of the droplet size distribution. From a qualitative data analysis it can be postulated that agglomeration must have an impact on the droplet size distribution. This is due to the fact that with increasing time the smallest size fraction disappears first, afterwards the 2^{nd} smallest and so on. This behaviour cannot be explained by evaporation alone. Furthermore, phenomena like sedimentation and convection or dispersion cannot cause this effect because sedimentation has a stronger effect on bigger droplets while the latter two are independent of droplet size. The disappearance of the smallest droplet size fraction cannot be explained by breakage because even smaller droplets are not detected. An analysis of the experimental data for the three biggest discrete droplet size fractions reveals that their percentage in the droplet size distribution starts decreasing at a certain time. This cannot be explained by evaporation because the evaporation does have a much weaker impact on the droplet size of larger droplets than of medium or small droplets. For this reason, either sedimentation or breakage or both must play a role. For the given time-scale all droplets that

sediment out of the measurement zone are replaced by sedimentation into the measurement zone from higher zones within the radius of the spraying circle detected during the experiments. The findings of the qualitative analysis can be summarized as follows:
Evaporation – decrease of droplet sizes and disappearance of droplet (needed)
Sedimentation – decrease of total number of droplets and increase of ratio of small droplets compared to big droplets (not needed)
Convection – transport of droplets in and out of domain, does not depend on droplet size (not considered)
Dispersion – transport of droplets in and out of domain, does not depend on droplet size (not considered)
Agglomeration – increase of droplet sizes and decrease of total number of droplets (needed)
Breakage – decrease of droplet sizes and increase of total number of droplets (might be needed)

Analysing the measured experimental data, three modelling scenarios (see Table 5.31) with increasing degree of detail are to be considered.

Table 5.31 *Multi-scale scenarios of interest*

Scenario	***Considered phenomena***
scenario 1	evaporation
scenario 2	evaporation + agglomeration
scenario 3	evaporation + agglomeration + breakage

Since it is good practice to start simple, the starting point is scenario 1 considering only evaporation. It is investigated to what extent this scenario is able to reflect the experimental data. It is sure that evaporation occurs to some extent and it is of interest that the model is able to predict the amount of limonene evaporating from the droplets with respect to time. Based on scenario 1 the degree of detail is gradually increased and it is analysed if and to what extent the model performance is improved in terms of the performed qualitative analysis.
In the following, the application of the multi-scale model construction work-flow is highlighted only for scenario 1.

Step II.1: Model-scenario documentation and concept (Phase II)

The documentation and concept for scenario 1 is briefly summarized in the Table 5.32.

Table 5.32 *Scenario documentation and concept*

Considered phenomena	***Evaporation***
Required models/scales:	
macro scale	behaviour of entire droplet population (corresponding population balance for scenario 1 with problem-specific discretization with respect to droplet size), size of measurement domain: 7.85 cm^3
micro scale	evaporation of single droplets, size of droplets 1.166-135.936 µm
properties	constant and temperature dependent properties required for evaporation model

Step II.2: Model development for new/current scale (Phase II)

Macro scale

The form of the generic population balance (Equation 5.304) corresponding to this scenario is given below:

$$\frac{\partial \underline{\Psi}(d,\underline{z},t)}{\partial t} = -\frac{\partial(\underline{\Psi} G)}{\partial d} \tag{5.316}$$

The droplets have been grouped into 31 discrete size fractions (see Phase I, Step I.2.6). Consequently, the discretization of the population balance results a system of 31 ODEs, one for each discrete diameter fraction.

Micro scale

The equations for the evaporation model are retrieved from the model library. Table 5.33 gives an overview of the assumptions and considered degree of detail of the applied evaporation model.

The corresponding model equations for the growth rate G_i of the droplets of size fraction i due to evaporation are given below. Since the droplet size decreases due to evaporation the growth rate G_i has a negative value.

$$R_{Kelvin,i} = exp\left\{\frac{-4 \cdot MW_i \cdot \sigma_i}{\rho_i \cdot Rm \cdot T_{s,i} \cdot S_i}\right\} \tag{5.317}$$

$$P_i^S = P_i^{s,plane} \cdot R_{Kelvin,i} \tag{5.318}$$

$$F_i = \frac{2 \cdot \pi \cdot D_{air} \cdot MW_i \cdot S_i \cdot P}{Rm \cdot T_{am}} ln\left\{\frac{1 - P_i^S/P}{1 - P_{am}/P}\right\} \tag{5.319}$$

$$G_i = F_i \cdot 2/\left(\pi \cdot \rho_i \cdot {S_i}^2\right) \tag{5.320}$$

$$0 = 2\pi d_i K_{air}\left(T_{am} - T_{s,i}\right) + \pi S_i^2 \Gamma\left(T_{am}^4 - T_{s,i}^4\right) + L_i F_i + cp_{vap,i} \cdot F_i \cdot T_{s,i} \tag{5.321}$$

Table 5.34 gives an overview of the required constants and temperature dependent properties and their sources.

Table 5.33 *Assumptions and considered phenomena for evaporation model*

Assumptions:	***Considered phenomena/details***
▪ Spherical droplets ▪ VLE established at droplet surface at all times ▪ Vapour behaves like ideal gas ▪ Steady state for energy balance of droplets ▪ Effect of other droplets on evaporation neglected ▪ Diffusivity through air of each vapour compound unaffected by presence of other vapour compounds ▪ Neglect Stefan flow ▪ Neglect convection effects ▪ Concentration of droplet compounds far from droplet is 0 ▪ Droplets are ideally mixed ▪ Constant diffusion coefficient (at Tam) ▪ Neglect heat and mass transfer resistance at droplet surface	▪ Diffusion of fragrance compound from droplet in gas phase ▪ Heat conduction from droplet in gas phase ▪ Radiation ▪ Gas kinetic transfer ▪ Kelvin effect (change of vapour pressure for droplet compared to vapour pressure over plane surface)

Table 5.34 *Properties required for droplet evaporation model and their sources*

properties	***sources***
temperature-dependent properties	
$cp_{vap,i}(T)$	ICAS-database (database is connected to generic computer-aided modelling framework)
$\rho_i(T)$	ICAS-database
$K_{air}(T)$	ICAS-database
$D_i(T)$	model library (model from (Wilke & Lee, 1955)
$\sigma_i(T)$	ICAS-database
$L_i(T)$	ICAS-database
$P_i^S(T)$	(Landoldt-Börnstein, 2011) (also available in model library and in ICAS database)
Constant properties	
MW	ICAS-database
T_c	ICAS-database

Step II.3: Establish data-flow scheme for new/current scale (Phase II)

The fragrance spraying template provides the data-flow scheme of this scenario (see Figure 5.31).

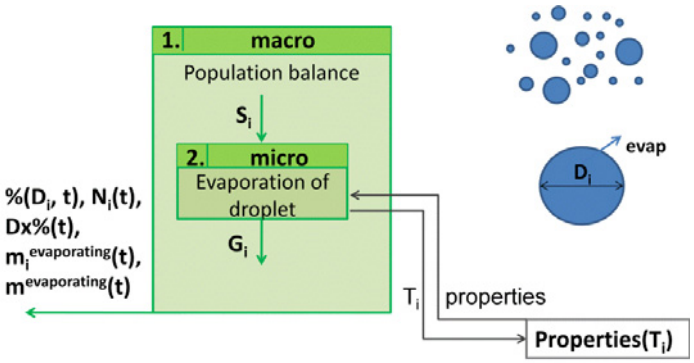


Figure 5.31 *Data-flow scheme between the scales of scenario 1.*

The macro scale needs to communicate the mean diameters S_i of the discrete size fractions of the droplets size distribution to the micro scale. On the micro scale the evaporation rate of the droplet of size S_i is calculated and communicated to the population balance model.

Step II.4: Derive linking scheme and link models accordingly (Phase II)

For the multi-scale models from Step II.3, a linking scheme is employed to obtain one overall model. Figure 5.32 shows the linking scheme for scenario 1.

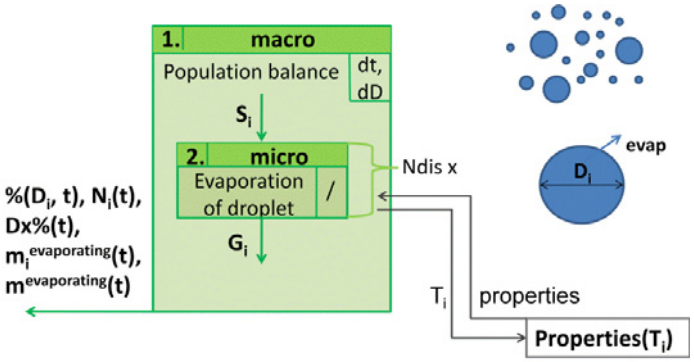


Figure 5.32 *Linking scheme for scenario 1 (phenomena considered: evaporation).*

The linking scheme shows the final data-flow between the scales as well as additional information like for example how often the different scales need to be solved.
For scenario 1 the micro-scale evaporation model needs to be solved $Ndis$ times, once for each discrete droplet size. The micro and macro scales do not need to be solved together but they may be solved sequentially because it is assumed that the droplet energy balance is at steady state from the beginning of the simulation and thus the droplet temperature and the evaporating mass stream do not change with time.

Step II.5: Overall model analysis (numerical) (Phase II)

Table 5.35 shows the number and types of the equations of the linked model.

Table 5.35 *Number and types of equations of model for scenario 1*

total number of equations	753	
number of ordinary differential equations (ODEs)	32	
number of algebraic equations	721	
	implicit	31
	explicit	690

The model variables can be pre-classified as 32 differential variables (1 ODE for each of the 31 discrete droplet size fractions between 1.166 µm and 135.936 µm resulting from the experiments and 1 ODE for the size fraction between 0 and 1.166 µm) and 745 algebraic variables. The algebraic variables consist of the variables of the evaporation models for the 31 discrete droplet size fractions (see Equations 5.317-5.321, e.g. droplet temperatures or vapour pressures) and the variables appearing in the ODE system resulting from the population balance (see Equation 5.316, e.g. evaporation rate for each discrete droplet size). To satisfy the degree of freedom, 24 algebraic variables need to be classified as either parameter or known variables. In scenario 1 all 24 variables are classified as known variables. Of the remaining 721 variables, 31 are classified as unknown implicit algebraic variables (these are the droplet temperatures of the droplets from the 31 measured different discrete droplet size fractions) while the remaining are classified explicit unknown variables. A simplified version of the resulting incidence matrix is given in Table 5.36. The incidence matrix is divided into three parts. The first two parts represent the property models and the evaporation models for the discrete droplet size fractions, respectively. The third part represents the population balance model. It can be noted that the first two parts can be solved independently from the last part while the population balance model needs results from the evaporation model. For this reason, the property models and the evaporation model are solved first for all 31 droplet sizes. These models need to be solved simultaneously due to the unknown droplet temperature (see off-diagonal element for property models in Table 5.36).

Initial guesses need to be provided for the droplet temperatures. Once the evaporation model has converged the resulting values for the growth rates G_i of the droplets are used to solve the population balance model. In order to do so, initial values for the concentrations of each droplet size fraction need to be provided. These values are used to evaluate the right hand side (RHS) of the ODEs based on which the droplet concentrations N_i for each time step is updated.

Table 5.36 *Simplified incidence matrix for scenario 1*

	Properties $(Ndis)$	$S_1(Ndis)$	$R_{Kelvin,i}(Ndis)$	$P_i^s(Ndis)$	$G_i(Ndis)$	$T_i(Ndis)$	$Dx\%(3)$	$Ntot$	$fr(Ndis)$	$dN_i(Ndis+1)$
Properties($Ndis$)	*					⊛				
$S_1(Ndis)$		*								
$R_{Kelvin,i}(Ndis)$	*	*	*			⊛				
$P_i^s(Ndis)$	*		*	*						
$G_i(Ndis)$	*	*		*	*	⊛				
$T_i(Ndis)$	*	*			*	*				
$Dx\%(3)$							*			⊛
$Ntot$								*		⊛
$fr(Ndis)$									*	⊛
$dN_i(Ndis+1)$		*			*					*

Model evaluation/ validation (Phase IV)

In order to evaluate the performance of scenario 1 the predictions of the 10%, 50% and 90% statistic diameters with respect to time have been plotted together with the corresponding experimental measurements (see Figure 5.33). The 10% statistic diameter (D10%), for example, is the diameter for which 10% of the droplets from the entire droplet size distribution have a lower diameter whereas for the D90%-diameter, 90% of the droplets have a lower diameter.

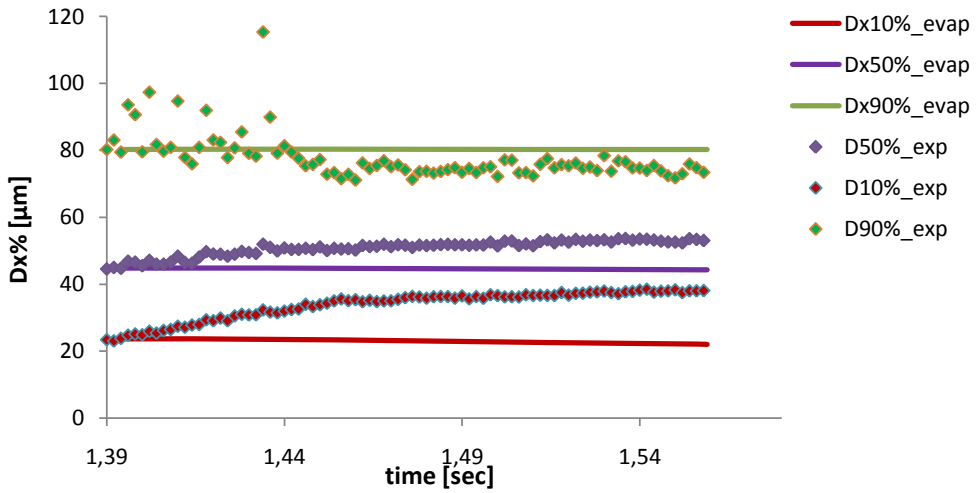


Figure 5.33 *Comparison of experimental data with model predictions for scenario 1 (evaporation): Statistic D10%, D50% and D90% diameters with respect to time.*

Figure 5.33 shows very clearly that the model is not able to follow the observed trend confirming that evaporation alone does not capture the behaviour of the system. In order to improve the model the following three options have been considered:

1. Improve values of the model parameters
2. Improve models for the included phenomena
3. Extend model by adding new phenomena from generic model (might add new scales)

In order to address the first two points a sensitivity analysis has been performed as it can give an indication if scenario 1 which only considers droplet evaporation is able to match the experimental data, for example, with different parameter values or a more detailed model for evaporation. During the sensitivity analysis the calculated evaporation rates have been multiplied by a constant factor. Different factors have been tested (0.1, 6, 20). Figure 5.34 shows the D10%-, D50%- and D90%-profiles for a factor of 6 in comparison to the previous profiles (factor 1) as well as the experimental profiles.

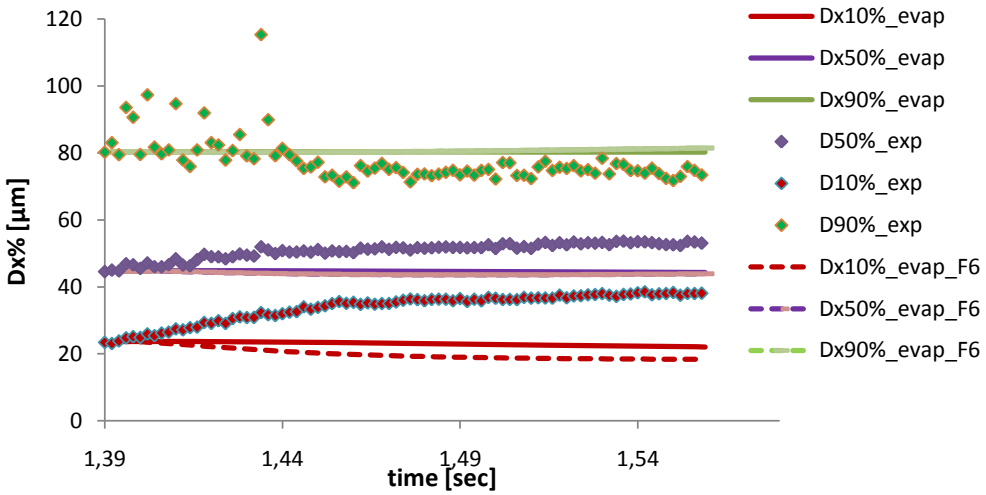


Figure 5.34 *Sensitivity analysis results for scenario 1 (increase of evaporation rate by factor of 6).*

The sensitivity analysis reveals that for none of the selected factors the behaviour of the system can be captured by the model (see Figure 5.34 as an example). This is, the system behaviour cannot be described by evaporation alone and more complex scenarios need to be investigated.

Evaluation/validation for scenario 2

Scenario 2 considers evaporation and agglomeration of droplets. The corresponding model has been developed following the steps of the multi-scale model construction work-flow combined with the fragrance spraying template similar to scenario 1. For scenario 2 and also for scenario 3

the different work-flow steps are not repeated in detail but instead just the main features of the model and the validation results are summarized.

The form of the population balance for scenario 2 is given by:

$$\frac{\partial N(d,\underline{z},t)}{\partial t} = -\frac{\partial (NG)}{\partial d} + R_A \tag{5.322}$$

The population balance is discretized with respect to droplet diameter according to the discrete size fractions given in Appendix A5 applying the method of classes. The agglomeration rate R_A for the different discrete droplet size fractions is determined applying the agglomeration model available in the model library (see Section 5.4.1).

The model for scenario 2 contains two unknown parameters C1 and C2 (see agglomeration model in Section 5.4.1) which result from the frequency and efficiency terms of the agglomeration rate equations. These two parameters need to be identified using the available experimental data; that is the dynamic %-profiles of the different discrete size fractions. The model identification has been performed applying the model identification work-flow provided by the generic computer-aided modelling framework together with the required tools and expert knowledge (Chapters 2 and 3). Only the main results are given below.

A local differential sensitivity analysis has been performed in order to identify if the available experimental data is sensitive to the parameters to be identified and to get an impression of the importance of the different parameters. The sensitivity analysis has not only been performed with respect to the unknown model parameters to be identified but also with respect to a constant factor by which the droplet evaporation rate has been multiplied. The base value for this factor was 1. The reason for the inclusion of the sensitivity of the evaporation rate is simply to get an impression on the significance of the evaporation impact compared to the agglomeration impact within the model. All parameters have been perturbed by ± 0.1%. The parameter significance ranking resulting from the analysis is shown in Figure 5.35. The parameters have been ranked according to the sensitivity measure $\delta_{msqr,j}$ which, for each parameter j, combines the sensitivities determined at all data points (for all measured variables at all measurement times) and is given in Equation 4.3.

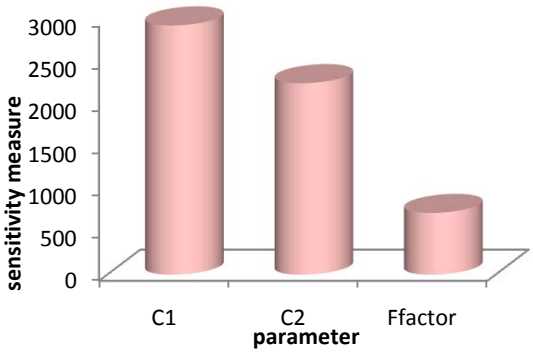


Figure 5.35 *Parameter significance ranking for agglomeration parameters C1 and C2 as well as evaporation rate factor (local differential sensitivity analysis at all data points, perturbation ± 0.1%).*

It can be detected from Figure 5.35 that the experimental data is sensitive with respect to all parameters but that the agglomeration parameters have a stronger impact.

An identifiability analysis has been performed (see Chapter 4) including all three parameters to identify the correlations between these parameters. There were no significant correlations between the parameters. Finally, the agglomeration rate parameters (C1, C2) have been estimated with the available experimental data. The parameter estimation itself has been performed using the SQP optimizer in MoT to minimize the following objective function:

$$Obj = \frac{1}{N_{dat}} \sum_{k=1}^{N_{dat}} (c_{\%,exp}(k) - c_{\%,sim}(k))^2 \tag{5.323}$$

The final value of the objective function was 1.295. The resulting value for C1 and C2 are 4.87×10^8 and 1.01×10^7, respectively. Figure 5.36 shows a comparison between the model predictions and the experimental measurements for the dynamic profiles of the D10%-, D50%- and D90%-diameters. Figure 5.36 shows a comparison between the model predictions of the dynamic droplet size distribution profiles and corresponding experimental data (this data has been used to identify the agglomeration rate parameters).

Scenario 2 is representing the experimental data of the system quite well and by far better than scenario 1. However, there are still some deviations for the profiles of the three biggest droplet size fractions. The percentage of these size fractions is predicted to slightly increase by the model whereas the experimental data shows a decrease (see Figure 5.36 bottom, right). For this reason, scenario 3, which considers evaporation, agglomeration and breakage is investigated

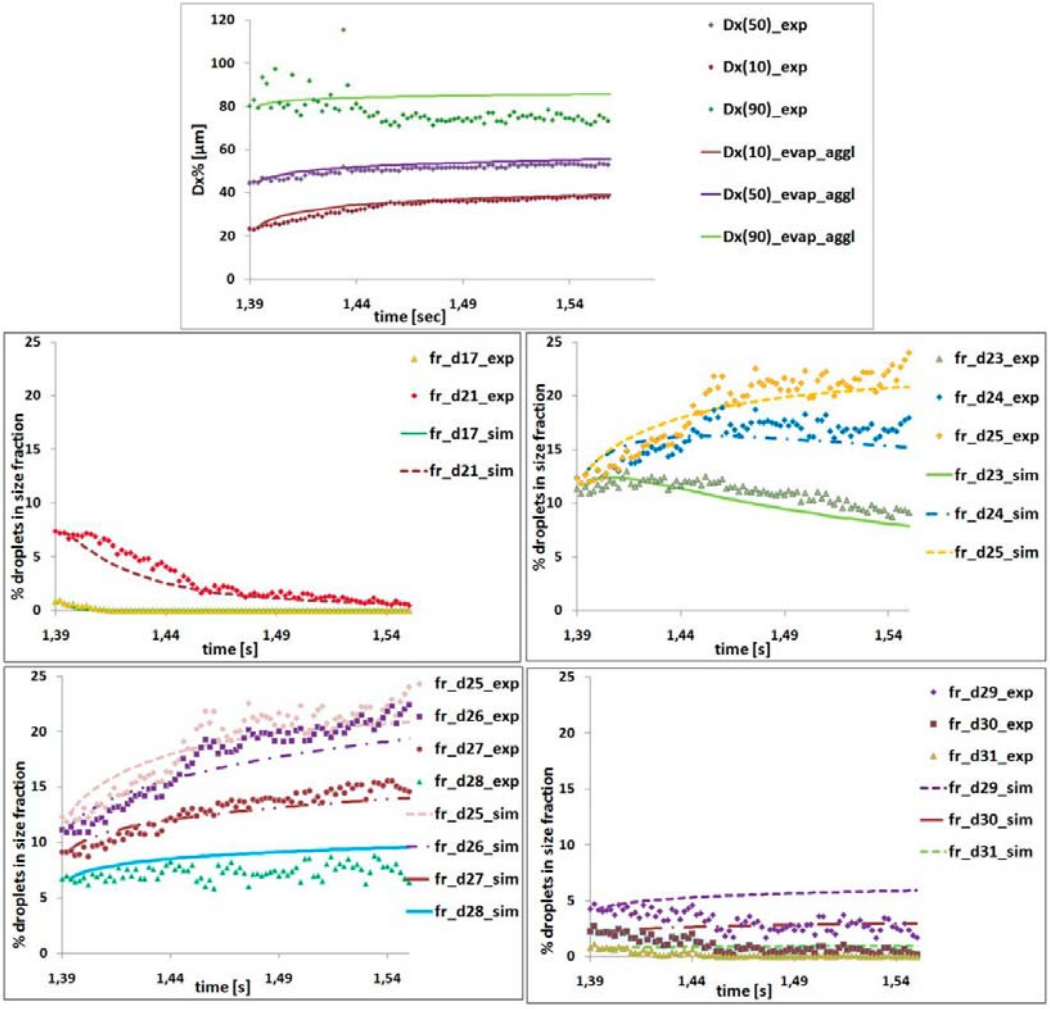


Figure 5.36 *Comparison of experimental data with model predictions for scenario 2 (evaporation and agglomeration): Statistic D10%-, D50%- and D90%-diameters (top) and percentage of different discrete size fractions in droplet size distribution (bottom) with respect to time.*

Evaluation/ validation of scenario 3

The population balance derived from the generic model for scenario 3 is given by:

$$\frac{\partial N(d,\underline{z},t)}{\partial t} = -\frac{\partial (NG)}{\partial d} + R_A + R_B \tag{5.324}$$

A constitutive model for breakage needs to be added, which is done by retrieving the breakage model described in Section 5.4.1 from the model library. Also for this scenario, parameter estimation is required. The three estimated parameters are C1 and C2 for the agglomeration rates as well as C3 for the breakage rates. The resulting value of the objective function

(Equation 5.321) results to be 0.497 for scenario 3 which is smaller than for scenario 2. The resulting parameter values for C1, C2 and C3 are 1.80×10^8, 6.92×10^6 and 2.49×10^3, respectively. Figure 5.37 shows the comparison of scenario 3 with estimated parameter values and the experimental data.

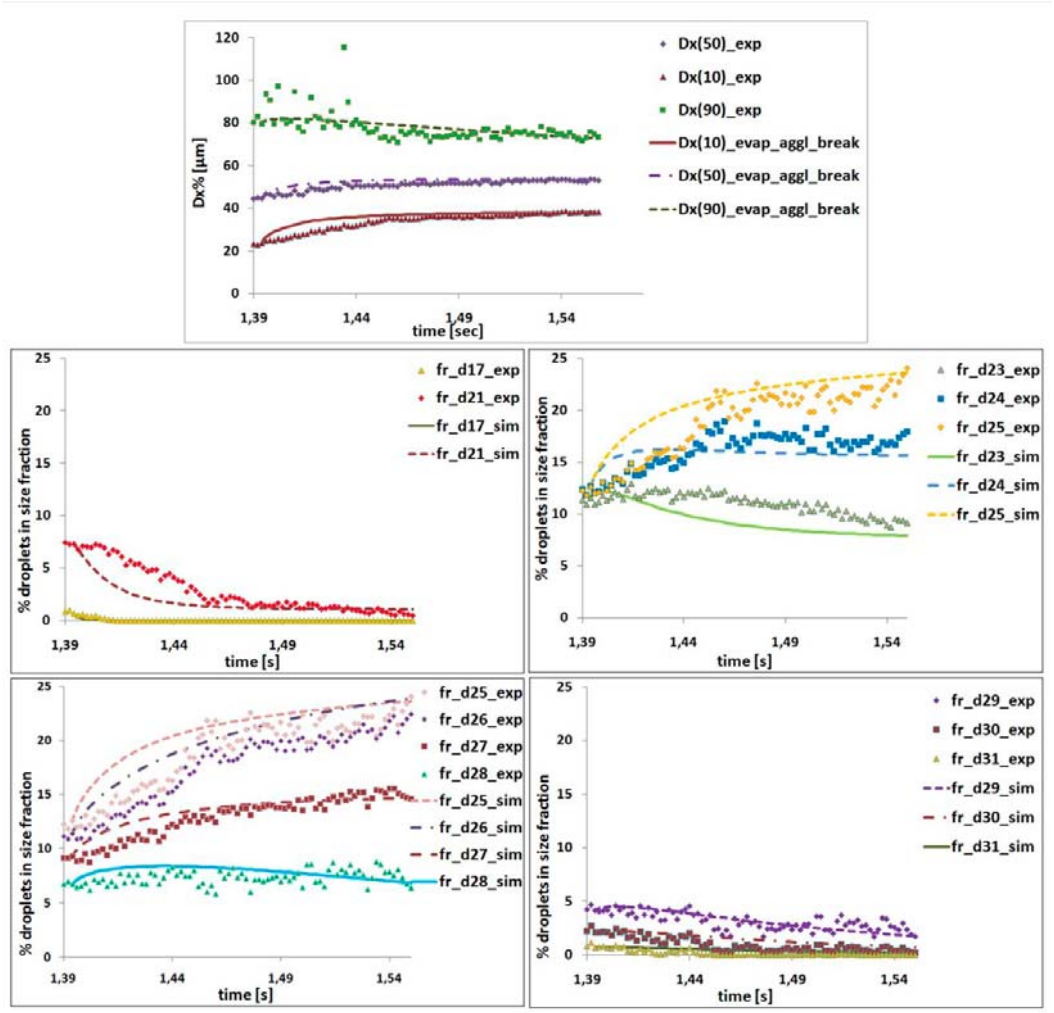


Figure 5.37 *Comparison of experimental data with model predictions for scenario 3 (evaporation, agglomeration and breakage): Statistic D10%-, D50%- and D90%-diameters (top) and percentage of different discrete size fractions in droplet size distribution (bottom) with respect to time.*

It can be seen that scenario 3 is representing the experimental data well and in contrast to scenario 2 is able to capture the decrease of percentage of the three biggest droplet size fractions at high times (see Figure 5.37 bottom, right).

In conclusion scenario 3 has been selected as the final model. The experimental data can be applied to identify (rate parameters for agglomeration and breakage models) and validate this special version of the generic model.

5.4.3 Fragrance aerosol case study 2

While the first case study is focused on deriving an appropriate model with respect to a modelling objective and available experimental data the second case study is application-oriented. It highlights how a developed fragrance aerosol model allows the evaluation of product attributes such as how much vapour and of which composition is released at which height and at what time; how fast droplets settle down and when they disappear due to evaporation. This allows the product designer to virtually experiment different scenarios (e.g. starting composition of fragrance) and design the appropriate product. For this case study droplet evaporation, sedimentation and dispersion in the vertical direction are considered. The corresponding models are included in the modelling framework library. Simulations have been conducted for a total number of 1.02×10^{10} droplets having a droplet size distribution of 22 discrete diameters between 1.3 and 34 µm. Initially, all droplets have a composition of 5 vol% limonene (active ingredient) and 95 vol% ethanol (solvent). Correlations for the pure component properties with respect to changing temperature are taken from the CAPEC database in ICAS which is connected to the modelling framework. For the activity coefficients with respect to temperature and concentration the UNIQUAC model has been applied. The UNIQUAC parameters have been fitted to experimental data for the system under consideration by Cháfer et al (2004). The micro-scale results are given in Figure 5.38. It can be detected that the distribution of droplets gets broader with increasing droplet diameter (Figure 5.38 right). This is due to the fact that the lifetime of bigger droplets is longer.

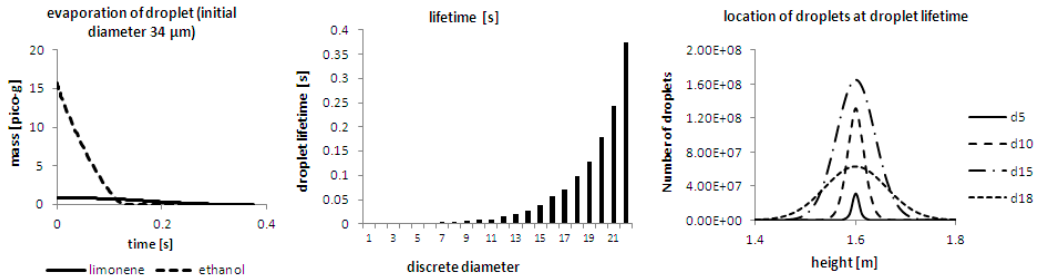


Figure 5.38 *Micro scale results. Left: droplet composition during evaporation (34 µm droplet), Centre: lifetimes of droplets for all 22 discrete diameters, Right: Location of droplets at droplet lifetime for different discrete diameters.*

Figure 5.39 shows the macro scale results of interest, that is, the total mass of limonene that has been released by all droplets at a given time (left) and how much ethanol (centre) and limonene (right) are released, in which height and at what time. This case study has been published in Heitzig et al. (2010).

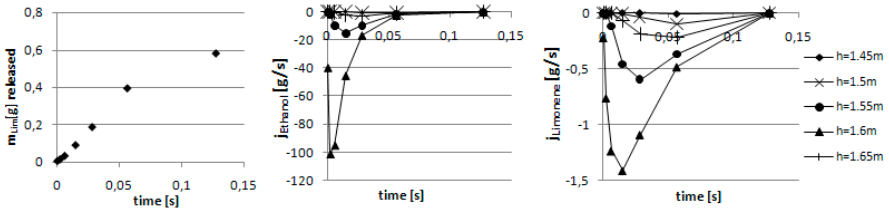


Figure 5.39 *Macro scale results. Left: Total mass of limonene released (by all droplets) vs. time, Right: Total mass flow of ethanol and limonene at different heights vs. time.*

5.4.4 Conclusions for fragrance aerosol template and case study

5.4.3.1 Fragrance spraying template

The fragrance spraying template has been developed as a system-specific template super-imposed to the generic work-flows of the computer-aided modelling framework. The template is one example of how problem-specific work-flows in general can be incorporated in the computer-aided modelling framework. The result is a combination of maximum flexibility and general support provided by the generic computer-aided modelling framework with problem-specific support added by the templates in one modelling framework.

The fragrance aerosol system has been chosen as an example because the fast and efficient development and application of reliable models with appropriate degree of detail to predict the behaviour of fragrance aerosols is a challenging problem of high interest for the related industries.

The main additions of the fragrance spraying template to the generic work-flow based framework are the provision of structured domain knowledge where needed, for example, the generic model with the corresponding model libraries, the speed-up of the model derivation process for different modelling goals for fragrance spraying systems and the strategy to derive the required properties for fragrance compounds.

5.4.3.2 Case studies

The first case study highlights the application and benefits of the fragrance spraying template in systematically deriving, analysing, identifying and validating a fragrance aerosol model

corresponding to a certain modelling objective. The systematically derived final model applying the proposed fragrance spraying template is able to reflect the experimental data obtained for the population balance from the performed dynamic light scattering experiments. Since agglomeration and breakage turn out to play an important role under the experimental conditions the data can be applied to identify the two parameters of the agglomeration rate model as well as the parameter of the breakage rate model. The generic model could be partly validated with the experimental data (agglomeration, breakage and evaporation phenomena).
The second case study demonstrates how the derived models can be applied for monitoring of product behaviour under differing conditions.

5.5 Pharma industry: Pharmacokinetic modelling of drug uptake and distribution in rats and humans

This case study demonstrates the application of the computer-aided modelling framework for a problem of high interest to the pharma-industry – the pharmacokinetic modelling of drug uptake and distribution in rats and the scale-up of the obtained model to humans. The goal is to develop a whole body physiologically-based pharmacokinetic (PBPK) model for the distribution of drugs in the body of rats based on a systematic engineering approach. The developed model is scaled-up to a human being. Traditionally, the modelling of pharmacokinetics is based on simple empirical curve-fitting of experimental data which has two main drawbacks:

1. Lack of physiological insight
2. Poor extrapolation quality (to different conditions, within individuals of the same species, between different species or for different drugs)

The engineering approach in contrast is based on conservation equations including transport and reaction terms and advanced scaling laws. The resulting models are mechanistic, based on the actual phenomena occurring in the body. The main advantages of this approach are:

- Increase of system knowledge: Gain of insights in the actual processes occurring during drug uptake and distribution;
- Better extrapolation ability, which has the potential to reduce the required number of undesirable as well as time- and cost-intensive animal experiments during drug development;
- Application of models for designing optimized and personalized drug dosage curves for patients and desired therapeutic effects.

However, the development of first principles PBPK models, their identification and discrimination between different candidate models is a non-trivial task inherent with a number of challenges that are related to the identification of the occurring phenomena and mechanisms within the body and especially finding the appropriate degree of detail of the models with respect to the modelling goal. The degree of detail is a trade-off between model complexity and model identifiability with accessible experimental data. In general PBPK models have a large number of equations and parameters. The developed modelling tool and its computer-aided work-flows provide a systematic strategy to address the problem of deriving, constructing and discriminating between different candidate models in order to find the model that is best supported by the experimental data available or accessible. Furthermore, the documentation and re-use of the different model candidates is enhanced. The benefit of the proposed systematic approach is an improvement of model quality and an increase of efficiency of the

modeller as well as the experimentalist and thereby a reduction of time and resources required for model development, identification, discrimination and application.

For the development of the pharmacokinetic rat model, as in the previous case studies, the work-flows for Phase I (Modelling objective and system information) and Phase II (Single- or multi-scale model construction) are highlighted. In Phase II the multi-scale model construction work-flow is applied to develop a multi-scale model where the different scales do not differ with respect to time or length scale but instead with respect to the degree of detail considered in the different models. Different candidate models are constructed (Phase II) and subsequently compared applying the model discrimination work-flow (Phase III). The pharmacokinetic modelling problem is the last case study presented. For this reason, not only the application of the different generic work-flows for this specific problem is illustrated here but also the support MoT provides to the modeller in the different work-flow steps is highlighted more clearly. The latter refers to the features that are available for the different work-flows and how the interface in MoT looks like. For the scale-up of the rat model to a human, the application of the MoT sensitivity and identifiability analysis features are highlighted. Here, the objective is to identify which model parameters are applicable for re-fitting with the limited experimental data (compared to the available rat data) available for humans.

5.5.1 Phase I. Modelling objective and system information

Step I.1: Modelling objective (Phase I)

The objective is to develop a model that is able to predict the distribution of the drug Cyclosporin A (CyA) in the body of a rat after inboulus injection. The dynamic drug concentration profiles in the major organs and in the blood are to be predicted.

The model needs to be a first principles based whole body model of the rat in order to allow extrapolation between different individuals, drugs and species. The model is to be applied to design optimal therapies and drug administration regimes.

Step I.2: System information and documentation (Phase I)

Step I.2.1: Functional description/ sketch of the system to be modelled (process, unit operation, product)

Figure 5.40 shows a sketch of the system to be modelled. The body of the rat is modelled as a network of organs connected by arterial (dotted line) and venous (solid line) blood vessels.

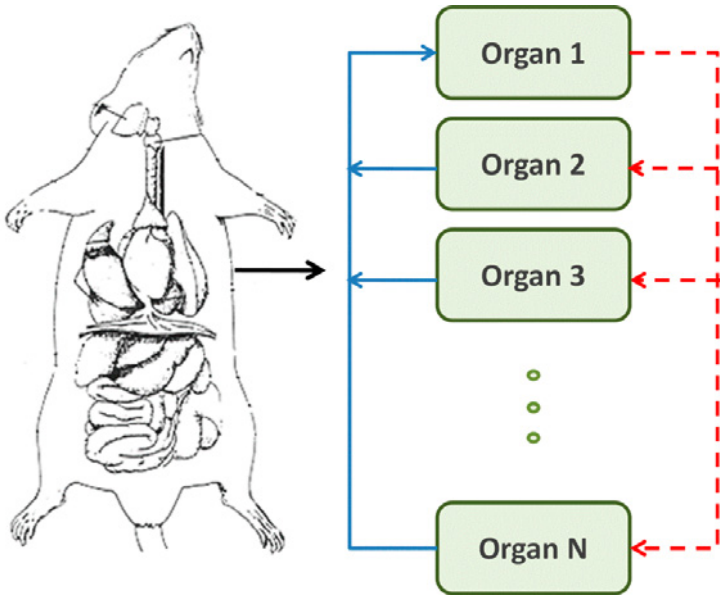


Figure 5.40 *Sketch of system to be modelled: Organ network of rat*

Which organs are to be specifically included in the model and the degree of detail of the different organ models as well as the metabolism need to be selected in a way for the model to satisfy the modelling objective. The degree of complexity should not be artificially high with respect to the modelling purpose because it has to be still supported by the structure and accuracy of the experimental data available.

Step I.2.2: System conditions

Step I.2.3: Phenomena in the system that might be of importance

The following phenomena might play a role for the system:

- Blood flow in main vessels (e.g. Hagen-Poiseuille)
- Blood flow in capillaries
- Multi-phase blood flow with transfer of drug between red blood cells (hct) and plasma
- Binding of drug to plasma proteins in blood
- Binding of drug to plasma proteins in interstitial fluid (IF)
- Mass transfer across cell membranes (can be also active transport)
- Mass transfer between interstitial fluid and cell tissue
- Metabolism of drug (reaction network)

Step I.2.4 Modelling of system/ problem

Whole body physiologically-based pharmacokinetic models for rates have been developed by a number of authors like Kawai et al. (1998), Mosat et al. (2011), Tanaka et al. (2000). For all these

models the body of the rat is modelled as a network of organs which are connected by different blood vessel segments. Each organ is related to an organ blood vessel segment, a venous segment and an arterial segment. Additional blood vessel segments are required to connect the blood flow between different organs in the network. The blood flow is assumed to consist of two phases – the plasma and the red blood cells (hct). The drug (CyA) is distributed in both phases. In plasma the drug occurs in two states. It can either be bound to plasma proteins (plB, plasma bound) or it can be free (plU, plasma unbound). Kawai & Lemaire (1993) have developed an equilibrium-correlation which relates the concentration of drug in total blood to its concentration in unbound plasma (plU) in equilibrium from in-vitro experiments. Lemaire & Tillement (1982) have derived a constant value (fuP=0.062) for the ratio of unbound drug concentration in plasma and total drug concentration in plasma. For the rat organs chamber models are applied, which for example, consider a separate chamber for the interstitial fluid (IF) and the cell tissue (CT). Mosat et al. (2011) propose an engineering approach to model the system considering two different scales. For both scales the network of organs and blood flow is considered but on two different degrees of detail. The low-detail-scale consists of a steady state blood flow model and does not consider the drug injection. The total blood flow (plasma and hct combined) is calculated for each blood segment (including organs). The resulting flow rates from the low-detail-scale are used in a more detailed model, which is a dynamic model and considers the drug injection, detailed organ models as well as multi-phase blood flow (hct, plasma, with binding of drug to plasma proteins).

Step I.2.5 Possible assumptions

Possible assumptions are:

- Steady state blood flow;
- Model blood flow by Hagen-Poiseuille;
- Compartment models for organs (assuming different compartments to be ideally mixed);
- Compartment model for blood vessel: hct, plasma bound, plasma unbound;
- Simplified metabolism;
- Metabolism of drug only in liver and kidney;
- Neglect some organs;
- Ratio of bound drug and unbound drug concentrations in plasma constant;
- Volume of compartments for unbound drug in plasma, bound drug in plasma and hct are constant;
- Blood capillaries are neglected (only consider main blood flow vessels combined with real mass transfer area of organs).

Step I.2.6 Preliminary system/process/reactor data

Experimental data is available from Tanaka et al. (2000). The authors have measured the total blood concentration of Cyclosporine A as well as the concentrations of the drug in 12 organs at different times in rats after a 2 minutes infusion of the drug via the femoral vein. Table 5.37 summarizes the conditions of the experiments.

Table 5.37 *Conditions of experiments (Tanaka et al., 2000).*

Mass or rats:	277 ± 15 g
Drug dose:	6 mg/kg rat in femoral vein
Injection time:	2 min
Replicate rats for each time point in dynamic organ concentration profiles:	3
Replicate rats for each time point in dynamic blood concentration profile:	at least 3 (6 for 30 min, 2h, 8h time points)

The organ concentrations at each time point were measured after killing the rat, taking out the different organs and homogenizing them. The blood concentration was measured by collecting blood from the jugular vein (except for 2 min and 30 min time groups an arterial catheter was used). Tanaka et al. (2000) state that the arterial-venous differences in the CyA blood concentrations are negligible, except for the first few minutes. In the experimental data (Tanaka et al., 2000) no replications but only the mean measurements and the standard deviations are given. Based on this information replications have been artificially created for this case study in a way that the published standard deviation and mean values are kept. The generated data based on Tanaka et al. (2000) is given in Appendix A6.

Step I.2.7 Select model-scenarios of interest

The two detail-scales are schematized in Figure 5.41. In the following the low-detail scale is also referred to as the macro-scale whereas the high-detail-scale is referred to as micro-scale.
From this generic scheme different scenarios can be derived by refining the detail of the organ and blood vessel models with respect to the modelling goal and corresponding experimental data. For this case study 12 organs are considered (lung, heart, brain, bone, fat, muscle, skin, liver, spleen, guts, kidney and thymus) because experimental measurements of the dynamic CyA concentration profiles are available for these organs. Possible modifications in the organ models are the number of compartments, transport mechanisms between compartments and the metabolism. Possible modifications of the blood vessel model are the transfer of drug between hct and plasma, the binding of drug to plasma proteins and the transport of drug to organs.

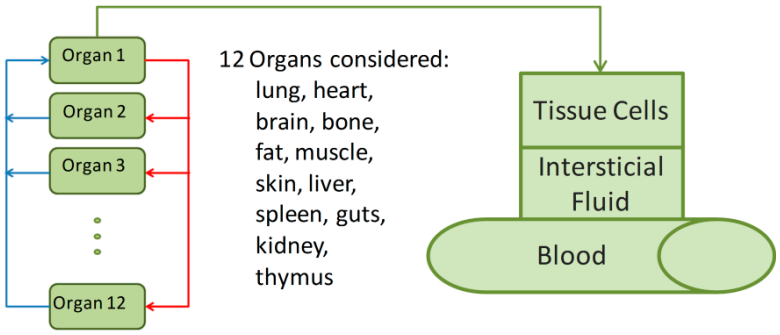


Figure 5.41 *General scheme for multi-scale scenarios*

In a first step the optimal blood vessel model is to be found.
Two candidates (see Figure 5.42) with alternative blood vessel models are developed in order to compare their performance in the later model discrimination phase.

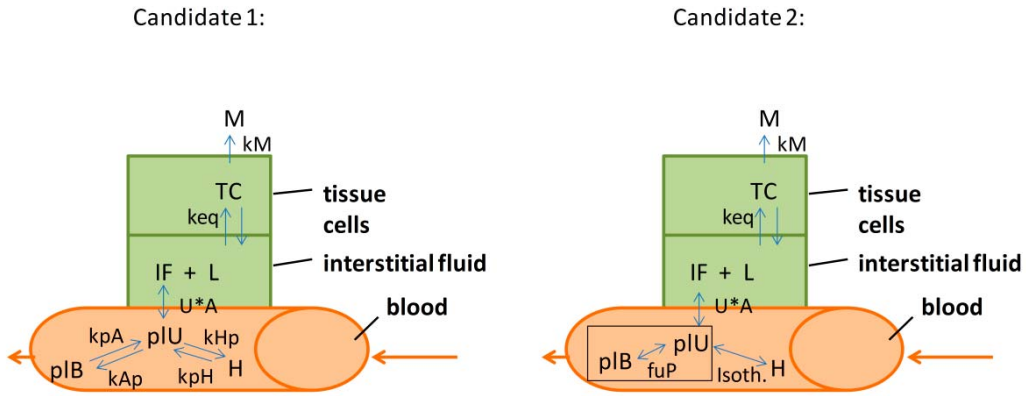


Figure 5.42 *Comparison between candidate 1 and 2.*

Both candidates have the same 2-compartment organ model (constitutive model) assuming a simplified metabolism represented by a 1st order reaction. The drug enters the interstitial fluid (IF) of the organ from the plasma unbound compartment (plU) of the blood vessels. The transport of CyA between plU and IF of the organs depends on the concentration difference in IF and plU whereas the transport of drug between interstitial fluid (IF) and tissue cells (TC) is assumed to be in equilibrium. The blood vessels are modelled to consist of three compartments: 1. plasma unbound (plU), 2. plasma bound (plB) and 3. red blood cells (hct). The concentration of CyA for the plasma bound compartment gives the concentration of CyA, which is bound to plasma proteins whereas the CyA concentration for the plU compartment gives the concentration of free CyA in plasma. The model-scenario for candidate 1 has been taken from Mosat et al. (2011) and Lueshen et al. (2011). Candidate 2 is a modified version from candidate 1, which in contrast to candidate 1 assumes equilibrium of the drug transport

between the blood vessel compartments and that the blood cell binding can reach saturation. That this is the case has been shown in vitro by Kawai & Lemaire (1993). They have also published an equilibrium-correlation which gives the plU-concentration of CyA as a function of the total blood concentration. Furthermore, the plasma protein binding is assumed to be linear. That this assumption is valid under certain conditions has been stated by Lemaire & Tillement (1982). The constant ratio between the plU and the total plasma concentration has been stated to be 0.062. Scenario 1 considers reversible 1st order 'reactions' for the drug transfer between the blood vessel compartments. Both scenarios are available in the MoT model library. In a later step, after having identified the optimal model for the blood vessels, different alternatives for the organ models will be discriminated.

The following features provided by MoT for the 'Modelling objective and documentation'-phase have been applied for this case study:

- Structured documentation interface (See Figure 4.7)
- Automated report generation (structured report from all information documented for different work-flow steps and sub-steps)
- MoT Model library which provides different PBPK models
- Experimental data-interface (provided data will be automatically used by MoT where needed, e.g. parameter estimation, and also included in the report)

5.5.2 Phase II. Model construction

The multi-scale model construction work-flow is selected. It needs to be followed in a loop for both multi-scale scenarios. Here, the model construction is shown in parallel for both multi-scale scenarios.

Step II.1: Scenario documentation and concept (Phase II)

The concepts for scenarios 1 and 2 are summarized in Table 5.38 and Table 5.39, respectively.

Table 5.38 *Concept for scenario 1 .*

assumptions:	▪ Steady state blood flow ▪ Model blood flow by Hagen-Poiseuille ▪ 2-Compartment models for organs: interstitial fluid and tissue cells ▪ Organ compartments ideally mixed ▪ Simplified metabolism: irreversible 1st order reaction from drug to metabolite ▪ Metabolism occurs in liver and kidney (with different reaction rates) ▪ Mass transfer of drug between IF and TC with ligands, mass transfer in equilibrium, ligand concentration incorporated in equilibrium constant ▪ Mass transfer resistance between unbound plasma proteins and interstitial fluid ▪ Consider only 12 organs (lung, heart, brain, bone, fat, muscle, skin, liver, spleen, guts, kidney, thymus) ▪ 3-Compartment model for blood vessel: hct, plasma bound, plasma unbound ▪ Volume of compartments for unbound drug in plasma, bound drug in plasma and hct are constant ▪ No saturation of blood cell binding (for the conditions present), modelled by 1st order reversible reactions
considered phenomena:	▪ Blood flow in main vessels (e.g. Hagen-Poiseuille) ▪ Multi-phase blood flow with dynamic transfer of drug between hct and plasma ▪ Dynamic binding of drug to plasma proteins in blood ▪ Mass transfer across cell membranes ▪ Mass transfer between interstitial fluid and cell tissue like a reversible reaction with a ligand (in equilibrium) ▪ Metabolism of drug (reaction network)
scales:	***Low degree of detail (macro scale):*** steady state blood flow network ***High degree of detail (micro scale):*** dynamic distribution of drug considering a network with detailed organ models, multiphase blood-flow and metabolism

Table 5.39 *Concept for scenario 2 .*

assumptions:	▪ Steady state blood flow ▪ Model blood flow by Hagen-Poiseuille ▪ 2-Compartment models for organs: interstitial fluid and tissue cells ▪ Organ compartments ideally mixed ▪ Simplified metabolism: irreversible 1st order reaction from drug to metabolite ▪ Metabolism occurs in liver and kidney (with different reaction rates) ▪ Mass transfer of drug between IF and TC with ligands, mass transfer in equilibrium, ligand concentration incorporated in equilibrium constant ▪ Mass transfer resistance between unbound plasma proteins and interstitial fluid ▪ Consider only 12 organs (lung, heart, brain, bone, fat, muscle, skin, liver, spleen, guts, kidney, thymus) ▪ 2-Compartment model for blood vessel: hct, total plasma ▪ Volume of compartments for total plasma and hct are constant ▪ Equilibrium of transfer of CyA between hct and plasma (blood cell binding) as well as plasma unbound and plasma bound (plasma protein binding) is established at all times ▪ Linear plasma protein binding: The concentration ratio of the unbound CyA concentration in plasma and the total CyA concentration in plasma is assumed to be constant at 0.062 (from Lemaire and Tillement, 1982). ▪ Non-linear blood cell distribution: The blood cell binding is saturable. The non-linear blood cell distribution (in equilibrium) is calculated applying the equilibrium-correlation determined in vitro by Kawai and Lemaire (1993).
considered phenomena:	▪ Blood flow in main vessels (e.g. Hagen-Poiseuille) ▪ Multi-phase blood flow with transfer of drug between hct and plasma (in equilibrium) ▪ Binding of drug to plasma proteins in blood (in equilibrium) ▪ Mass transfer across cell membranes ▪ Mass transfer between interstitial fluid and cell tissue like a reversible reaction with a ligand (in equilibrium) ▪ Metabolism of drug (reaction network)
scales:	***Low detail (macro scale):*** steady state blood flow network ***High detail (micro scale):*** dynamic distribution of drug considering a network with detailed organ models, multiphase blood-flow and metabolism

Step II.2: Model development for new/current scale (Phase II)

In order to construct the models for each scale, the single-scale model construction work-flow is applied. The application of the single-scale model construction work-flow has been highlighted in detail for the previous case studies. Consequently, the model construction step is not highlighted here. In the following the model equations as well as a summary of the results from

the numerical model analysis are given for the high-detail-scale model for both scenarios (1 and 2).

Figure 5.43 shows a screen shot of the translated model for scenario 1 in MoT. The model equations for the high-detail-scale for scenario 1 are given by:

Drug injection rate:

$$finj = \begin{cases} if\ 0 \le t \le 2\ min\text{: } \frac{m_{inj} \cdot m_{rat}}{t_{inj} \cdot V_b(47) \cdot fup} \\ \quad if\ t > 2\ min\text{: } 0 \end{cases} \tag{5.325}$$

Drug concentrations in unbound plasma of blood vessel compartment i:

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\left(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\right) - kph \cdot c_p(i) + khp \cdot c_h(i) \cdot \frac{fuh}{fup} - kpa \cdot c_p(i) + kap \cdot c_a(i)\frac{fua}{fup}, \quad i = 1, 14 - 46, 48 - 57 \tag{5.326-5.369}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\left(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\right) - kph \cdot c_p(i) + khp \cdot \frac{fuh}{fup} \cdot c_h(i) - kpa \cdot c_p(i) + kap \cdot \frac{fua}{fup} \cdot c_a(i) - kpQ(i) \cdot \frac{1}{V_b(i) \cdot fup} \cdot c_p(i) + kQp(i) \cdot \frac{1}{V_b(i) \cdot fup} \cdot c_Q(i), i = 2 - 13 \tag{5.370-5.381}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\left(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\right) - kph \cdot c_p(i) + khp \cdot c_h(i) \cdot \frac{fuh}{fup} - kpa \cdot c_p(i) + kap \cdot c_a(i)\frac{fua}{fup} + finj, \quad i = 47 \tag{5.382}$$

Drug concentrations in bound plasma of blood vessel compartment i:

$$\frac{dc_a(i)}{dt} = \frac{fua}{V_b(i)}\left(fl_{in}(i) \cdot c_{a,in}(i) - fl_{out}(i) \cdot c_a(i)\right) + kpa \cdot \frac{fup}{fua} \cdot c_p(i) - kap \cdot c_a(i), i = 1 - 57 \tag{5.383-5.439}$$

Drug concentrations in red blood cells of blood vessel compartment i:

$$\frac{dc_h(i)}{dt} = \frac{fuh}{V_b(i)}\left(fl_{in}(i) \cdot c_{h,in}(i) - fl_{out}(i) \cdot c_h(i)\right) + kph \cdot \frac{fup}{fuh} \cdot c_p(i) - khp \cdot c_h(i), i = 1 - 57 \tag{5.440-5.496}$$

Drug concentrations in organ j:

$$\frac{dc_Q(j)}{dt} = kpQ(j) \cdot \frac{1}{V_Q(j)} \cdot c_p(j) - kQp(j) \cdot \frac{1}{V_Q(j)} \cdot c_Q(j) - ke(j-1) \cdot c_Q(j), \ j = 2, 3 \tag{5.497, 5.498}$$

$$\frac{dc_Q(j)}{dt} = kpQ(j) \cdot \frac{1}{V_Q(j)} \cdot c_p(j) - kQp(j) \cdot \frac{1}{V_Q(j)} \cdot c_Q(j), \quad j = 4 - 13 \tag{5.499-5.508}$$

Metabolite concentrations in organ j:

$$\frac{dc_m(l)}{dt} = ke(l) \cdot c_Q(l+1), \quad l = 1, 2 \tag{5. 509, 5. 510}$$

Total blood volume in body:

$$V_{blood} = \sum_{i=1}^{57} V_b(i) \tag{5.511}$$

Overall drug concentrations in unbound plasma, red blood cells, bound plasma, total plasma and total blood in body:

$$c_{plasmaU} = \frac{fup}{(fup+fua)V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_p(i) \tag{5.512}$$

$$c_{hct} = \frac{1}{V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_h(i) \tag{5.513}$$

$$c_{plasmaB} = \frac{fua}{(fup+fua)V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_a(i) \tag{5.514}$$

$$c_{plasmaT} = c_{plasmaB} + c_{plasmaU} \tag{5.515}$$

$$c_{blood\_tot} = \left(fuh \cdot c_{hct} + (fua + fup) \cdot c_{plasmaT}\right) \tag{5.516}$$

Total mass of drug in body:

$$m_{inVivo} = fup \cdot c_{plasmaU} \cdot V_{blood} + fuh \cdot c_{hct} \cdot V_{blood} + fua \cdot c_{plasmaB} \cdot V_{blood} + \sum_{j=2}^{13} V_Q(j) \cdot c_Q(j) \tag{5.517}$$

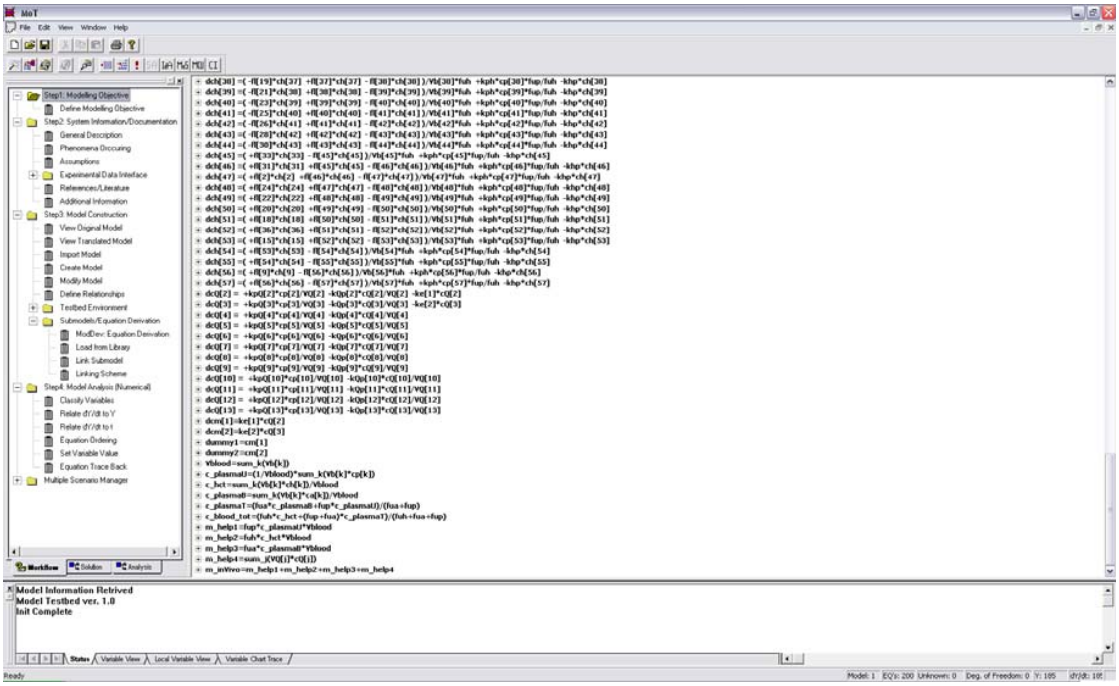


Figure 5.43 *Translated model in MoT (for scenario 1).*

The main results from the numerical model analysis for the high-detail-scale of scenario 1 are summarized in the following. The model consists of 8 explicit algebraic equations and 185 ordinary differential equations. The model variables can be pre-classified as 185 dependent variables ($c_p(i)$, $c_a(i)$, $c_h(i)$, $c_Q(j)$, $c_m(l)$; i=1-57, j=2-13, l=1,2), 1 independent variable (t) and 170 algebraic variables ($finj$, m_{inj}, m_{rat}, t_{inj}, fup, fua, fuh, kph, khp, kpa, kap, V_{blood},

$c_{plasmaU}$, c_{hct}, $c_{plasmaB}$, $c_{plasmaT}$, c_{blood_tot}, m_{inVivo},$V_b(i)$, $fl(i)$, $kpQ(j)$, $kQp(j)$, $V_Q(j)$, $ke(l)$; i=1-57, j=2-13, l=1,2). The degree of freedom analysis for the AE-part of the model results that 75 of the algebraic variables that appear in the AE-part need to be classified as either parameter or known. In this case all 75 variables are specified as known (m_{inj}, m_{rat}, t_{inj}, fup, fua, fuh, $V_b(i)$, $V_Q(j)$). The degree of freedom of the ODE-part corresponds to the number of algebraic variables that only appear in the ODE-part of the model. From these 87 variables, 30 are classified as parameters (kph, khp, kpa, kap, $pQ(j)$, $kQp(j)$, $ke(l)$) and the remaining 57 variables are classified as known variables ($fl(i)$). Table 5.40 gives the values of the known model variables for scenario 1 including their sources. The incidence matrix for the model is provided in Table 5.41.

Table 5.40 *Sources of known variables for micro-scale model, scenario 1*

known variable	***value***	***source***	***comment***
m_{inj}	6 mg/kg rat	Tanaka et al., 2000	corresponds to conditions of experimental data
m_{rat}	251.36 g	Tanaka et al., 2000	corresponds to conditions of experimental data, less than 277 g because not all organs considered in network
t_{inj}	2 min	Tanaka et al., 2000	corresponds to conditions of experimental data
$V_b(i)$	See MoT-file	Mosat et al., 2011	
$V_Q(j)$	See MoT-file	Kawai et al., 1998	
fup	0.37	Mosat et al., 2011	volume fraction of compartment
fua	0.18	Mosat et al., 2011	volume fraction of compartment
fuh	0.45	Mosat et al., 2011	volume fraction of compartment
$fl(i)$	See mot-file	macro scale model	

The initial conditions for the dependent variables are given by:

$c_p(i)(t=0)=0\,, \quad i=1-57$

$c_a(i)(t=0)=0\,, \quad i=1-57$

$c_h(i)(t=0)=0\,, \quad i=1-57$

$c_Q(j)(t=0)=0\,, \quad j=2-13$

$c_m(l)(t=0)=0\,, \quad l=1-2$

Table 5.41 *Incidence matrix for micro-scale model, scenario 1*

eq. no.	variable:	$finj$	V_{blood}	$c_{plasmaU}$	c_{hct}	$c_{plasmaB}$	$c_{plasmaT}$	c_{blood_tot}	m_{inVivo}	$c_p(i)$	$c_a(i)$	$c_h(i)$	$c_Q(i)$	$c_m(i)$
5.323	$finj$	*												
5.509	V_{blood}		*											
5.510	$c_{plasmaU}$		*	*						*				
5.511	c_{hct}		*		*							*		
5.512	$c_{plasmaB}$		*			*					*			
5.513	$c_{plasmaT}$			*		*	*							
5.514	c_{blood_tot}				*		*	*						
5.515	m_{inVivo}		*	*	*	*			*					
5.324-5.380	$c_p(i)$	*								*	*	*		
5.381-5.437	$c_a(i)$									*	*			
5.438-5.494	$c_h(i)$									*		*		
5.495-5.506	$c_Q(i)$									*			*	
5. 507, 5. 508	$c_m(i)$												*	*

The incidence matrix consists of an AE-part and an ODE-part. The latter is highlighted dark in Table 5.41. The AE-part can be transformed to a lower tridiangular form but it is coupled to the ODE-part of the model (see off-diagonal elements). The ODEs for c_p, c_a and c_h are coupled because they cannot be transformed to a lower tridiangular form (see off-diagonal elements). Furthermore, in order to update the RHS of the ODEs in each time-step results from the AE-part are needed (value for $finj$). Figure 5.44 shows a screen shot of the numerical model analysis in MoT for scenario 1.

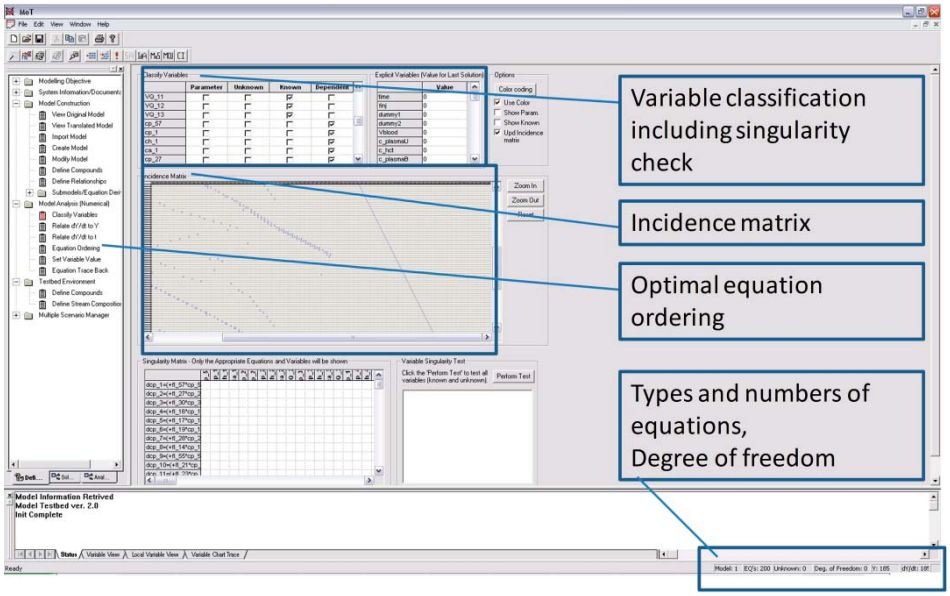


Figure 5.44 *Numerical model analysis in MoT (for scenario 1).*

The model equations for scenario 2 are as follows:

$$finj = \begin{cases} if\ 0 \le t \le 2\ min\text{:}\ \frac{m_{inj} \cdot m_{rat}}{t_{inj} \cdot V_b(47)} \\ if\ t > 2\ min\text{:}\ 0 \end{cases} \quad (5.518)$$

$$c_p(i) = 0.5 \cdot \left[c_b(i) - \left((1-fuh)\cdot fuP + fuh\right)\cdot K_D - fuh \cdot nP_T + \left(\left(c_b(i) - \left((1-fuh)\cdot fuP + fuh\right)\cdot K_D - fuh \cdot nP_T \right)^2 + 4 \cdot \left((1-fuh)\cdot fuP + fuh\right)\cdot K_D \cdot c_b(i) \right)^{\frac{1}{2}} \right] / \left((1-fuh)\cdot fuP + fuh\right) \quad (5.519\text{-}5.574)$$

$$c_{p\_tot}(i) = c_p(i)/fuP \quad (5.576\text{-}5.633)$$

$$c_a(i) = (1-fuP)c_{p\_tot}(i) \quad (5.633\text{-}5.689)$$

$$c_h(i) = \frac{1}{fuh} \cdot \left[c_b(i) - (1-fuh)\cdot c_{p\_tot}(i) \right] \quad (5.690\text{-}5.746)$$

$$\frac{dc_b(i)}{dt} = \frac{1}{V_b(i)} \left(fl_{in}(i)\cdot c_{b,in}(i) - fl_{out}(i)\cdot c_b(i) \right),\ i = 1, 14-46, 48-57 \quad (5.747\text{-}5.790)$$

$$\frac{dc_b(i)}{dt} = \frac{1}{V_b(i)} \left(fl_{in}(i)\cdot c_{b,in}(i) - fl_{out}(i)\cdot c_b(i) \right) - kpQ(i)\cdot \frac{1}{V_b(i)} \cdot c_p(i) + kQp(i) \cdot \frac{1}{V_b(i)} \cdot c_Q(i),\ i = 2-13 \quad (5.791\text{-}5.802)$$

$$\frac{dc_b(i)}{dt} = \frac{1}{V_b(i)} \left(fl_{in}(i)\cdot c_{b,in}(i) - fl_{out}(i)\cdot c_b(i) \right) + finj\,, \quad i = 47 \quad (5.803)$$

$$\frac{dc_Q(j)}{dt} = kpQ(j)\cdot \frac{1}{V_Q(j)} \cdot c_p(j) - kQp(j)\cdot \frac{1}{V_Q(j)} \cdot c_Q(j) - ke(j-1)\cdot c_Q(j)\,, \quad j = 2, 3 \quad (5.804, 5.805)$$

$$\frac{dc_Q(j)}{dt} = kpQ(j)\cdot \frac{1}{V_Q(j)} \cdot c_p(j) - kQp(j)\cdot \frac{1}{V_Q(j)} \cdot c_Q(j)\,, \quad j = 4-13 \quad (5.806\text{-}5.815)$$

$$\frac{dc_m(l)}{dt} = ke(l)\cdot c_Q(l+1)\,, \quad l = 1, 2 \quad (5.816, 5.817)$$

$$V_{blood} = \sum_{i=1}^{57} V_b(i) \quad (5.818)$$

$$c_{plasmaU} = \frac{1}{V_{blood}} \sum_{i=1}^{57} V_b(i)\cdot c_p(i) \quad (5.819)$$

$$c_{hct} = \frac{1}{V_{blood}} \sum_{i=1}^{57} V_b(i)\cdot c_h(i) \quad (5.820)$$

$$c_{plasmaB} = \frac{1}{V_{blood}} \sum_{i=1}^{57} V_b(i)\cdot c_a(i) \quad (5.821)$$

$$c_{plasmaT} = \left(c_{plasmaB} + c_{plasmaU} \right) \quad (5.822)$$

$$c_{blood\_tot} = \left(fuh \cdot c_{hct} + (1-fuh)\cdot c_{plasmaT} \right) \quad (5.823)$$

$$m_{inVivo} = (1-fuh)\cdot c_{plasmaT} \cdot V_{blood} + fuh \cdot c_{hct} \cdot V_{blood} + + \sum_{j=2}^{13} V_Q(j)\cdot c_Q(j) \quad (5.824)$$

The model consists of 71 ODEs and 236 (explicit) AEs. The model variables can be pre-classified in 71 dependent variables ($c_b(i)$, $c_Q(j)$, $c_m(l)$; i=1-57, j=2-13, l=1,2), 1 independent variable (t) and 395 algebraic variables ($finj$, m_{inj}, m_{rat}, t_{inj}, fuP, fuh, K_D, nP_T, V_{blood}, $c_{plasmaU}$, c_{hct}, $c_{plasmaB}$, $c_{plasmaT}$, c_{blood_tot}, m_{inVivo}, $V_b(i)$, $fl(i)$, $c_p(i)$, $c_a(i)$, $c_h(i)$, $c_{p\_tot}(i)$, $kpQ(j)$,

$kQp(j)$, $V_Q(j)$, $ke(l)$; i=1-57, j=2-13, l=1,2). The degree of freedom for the AE-part is 76. 73 variables are specified as known (m_{inj}, m_{rat}, t_{inj}, fuh, $V_b(i)$, $V_Q(j)$) and 3 variables are specified as parameters (K_D, nP_T, fuP).The degree of freedom for the ODE-part is 83. For this case 26 variables are specified as parameters ($pQ(j)$, $kQp(j)$, $ke(l)$) and the remaining 57 variables are specified as known ($fl(i)$). Table 5.42 gives the incidence matrix of the ordered equations.

Table 5.42 *Incidence matrix for micro-scale model, scenario 2*

eq. no.	variable:	$finj$	V_{blood}	$c_{plasmaU}$	c_{hct}	$c_{plasmaB}$	$c_{plasmaT}$	c_{blood_tot}	m_{inVivo}	$c_p(i)$	$c_{p\_tot}(i)$	$c_a(i)$	$c_h(i)$	$c_Q(i)$	$c_b(i)$	$c_m(i)$
5.323	$finj$	*														
5.509	V_{blood}		*													
5.510	$c_{plasmaU}$		*	*						*						
5.511	c_{hct}		*		*								*			
5.512	$c_{plasmaB}$		*			*						*				
5.513	$c_{plasmaT}$			*		*	*									
5.514	c_{blood_tot}				*		*	*								
5.515	m_{inVivo}		*		*		*		*					*		
5.324-5.380	$c_p(i)$									*					*	
	$c_{p\_tot}(i)$									*	*					
5.381-5.437	$c_a(i)$										*	*				
5.438-5.494	$c_h(i)$										*		*		*	
5.495-5.506	$c_Q(i)$									*				*		
5.438-5.495	$c_b(i)$	*								*				*	*	
5. 507, 5. 508	$c_m(i)$													*		*

The incidence matrix consists of an AE-part and an ODE-part (highlighted dark). In contrast to scenario 1 the ODEs are uncoupled for scenario 2 and can be brought to a lower tridiangular form. Also the AE-part is uncoupled. The AE-part and ODE-part are coupled with each other (see off-diagonal elements of AE-part). Table 5.43 gives an overview from which sources the known model variables are obtained.

Table 5.43 *Sources of known variables for micro-scale model, scenario 2*

known variable	*value*	*source*	*comment*
m_{inj}	6 mg/kg rat	Tanaka et al., 2000	Corresponds to conditions of experimental data
m_{rat}	251.36 g	Tanaka et al., 2000	Corresponds to conditions of experimental data, less than 277 g because not all organs considered in network
t_{inj}	2 min	Tanaka et al., 2000	Corresponds to conditions of experimental data
$V_b(i)$	See MoT-file	Mosat et al., 2011	
$V_Q(j)$	See MoT-file	Kawai et al., 1998	
fuP	0.062	Lemaire and Tillement, 1982; Tanaka et al., 2000	Concentration fraction of unbound CyA in plasma
K_D	0.185 µg/mL	Kawai and Lemaire, 1993; Tanaka et al., 2000	Dissociation constant (in equilibrium-correlation)
nP_T	4.64 µg Eq/mL	Kawai and Lemaire, 1993; Tanaka et al., 2000	Binding capacity (in equilibrium-correlation)
fuh	0.45	Mosat et al., 2011	Volume fraction of compartment (red blood cells)
$fl(i)$	See mot-file	Macro scale model	

The initial conditions for the dependent variables for scenario 2 are given below:

$c_b(i)(t=0)=0\,,\quad i=1-57$

$c_Q(j)(t=0)=0\,,\quad j=2-13$

$c_m(l)(t=0)=0\,,\quad l=1-2$

Figure 5.45 summarizes the MoT features used for this case study in the different steps of the single-scale model construction for the scale models.

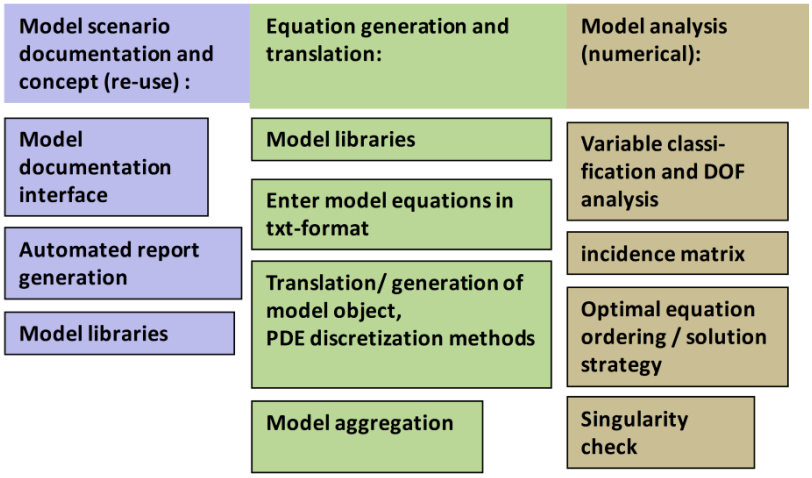


Figure 5.45 *Features used for single-scale model construction of the models for the different scales involved in scenarios 1 and 2.*

Step II.3: Establish data-flow scheme for new/current scale (Phase II)

The data-flow between the scales is shown in the linking scheme (Figure 5.46).

Step II.4: Derive linking scheme and link models accordingly (Phase II)

Figure 5.46 shows the linking scheme which is identical for both scenarios.

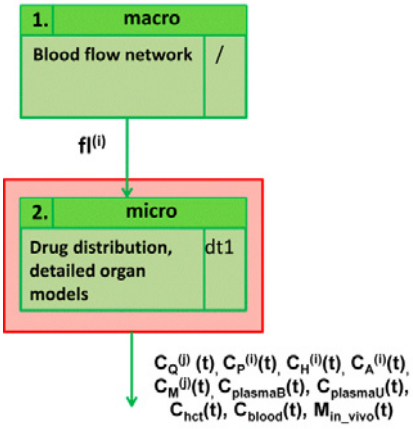


Figure 5.46 *Linking scheme for scenarios 1 and 2.*

The data-flow reveals that the two degree of detail scales can be solved sequentially. First, the steady state blood flow network model is solved and the resulting blood flows are communicated to the more detailed drug distribution model.

5.5.3 Phase III. Model identification/ discrimination

Step III.1: Experimental data (Phase III)

The experimental data available from Tanaka et al. (2000) (see Appendix A6) has been added to the MoT-data interface during Phase I and is automatically made available where needed during this work-flow.

Step III.2: Sensitivity analysis (Phase III)

Based on the experimental data provided and the variable classification (parameters) MoT automatically sets up the local differential sensitivity analysis. The resulting overall parameter significance ranking for scenario 1 based on the sensitivity measure given in Equation 4.3 with respect to all 30 model parameters at all available data-points is given in Figure 5.47.

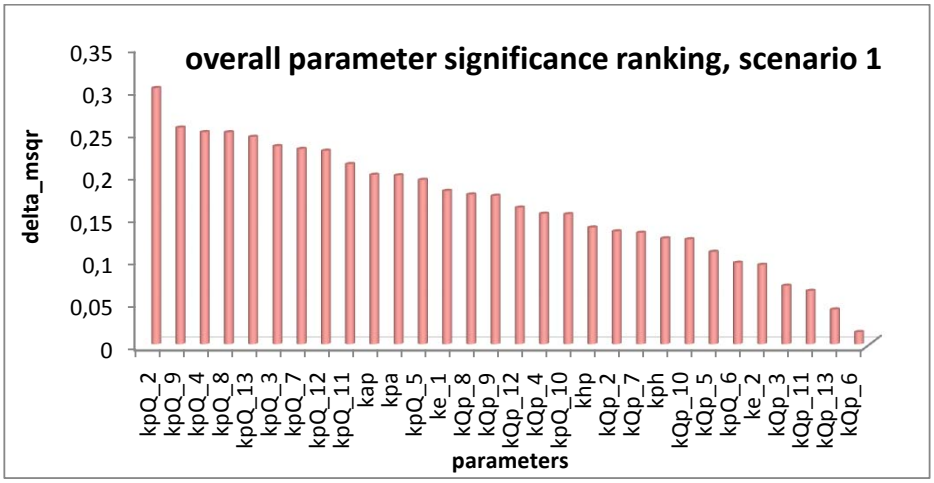


Figure 5.47 *Parameter significance ranking for scenario 1 (30 parameters).*

The parameter significance ranking for scenario 1 reveals that the only parameter that might be considered as not sensitive and therefore excluded from the further analysis is kQp6. However, this parameter has been considered in the following identifiability analysis.

Figure 5.48 shows the parameter significance ranking for scenario 2. Here, the situation is slightly different because fuP is by far the most sensitive parameter. However, the sensitivity measure of the remaining model parameters is not low enough to be excluded them from the further analysis.

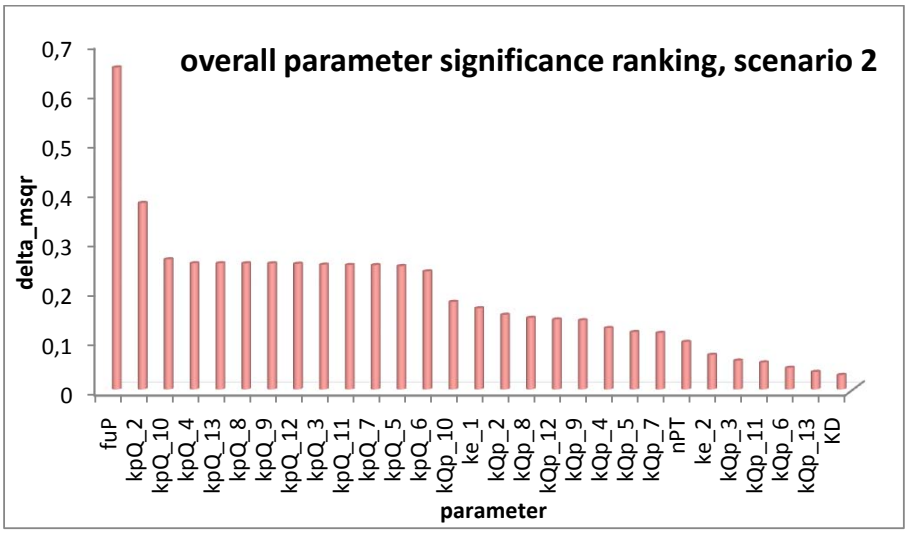


Figure 5.48 *Parameter significance ranking for scenario 2 (29 parameters).*

Step III.3: Identifiability analysis (Phase III)

The identifiability analysis is set-up automatically by MoT (see Chapter 4). The modeller is asked to select a threshold for the collinearity index and the parameters that are to be excluded from the analysis based on the sensitivity analysis results.

For scenario 1 the largest identifiable subset if found to be 28, that is, 2 parameters lower than the total number of parameters. The largest identifiable subset for scenario 2 is 2 parameters lower than the total number of parameters (27 parameters).

Step III.4: Parameter estimation (Phase III)

The parameter bounds have been set to 0 and $2x10^8$ for both scenarios whereas the initial parameter values are taken from Mosat et al. (2011). The objective function has been selected according to the model discrimination method developed by Stewart et al. (1992, 1996, 1998). The method is briefly summarized in Appendix A7.

For this case study the relevant case is 2-b, 'multiple measured variables, unknown covariance matrix' (Appendix A7) because the available experimental data consists of multiple measured variables and the covariance matrix of the measurement errors is unknown. However, the required replications of the measurements are available (generated artificially, see Phase I). Consequently, the objective function for the parameter estimation is the determinant $\left|\hat{v}_j\right|$ of the experimental error moment matrix $v_{ik}(\boldsymbol{\theta_j})$ (see Appendix A7, Equation A7.8). MoT automatically creates the optimization problem by transforming the model equations,

parameter boundaries and variable constraints to equality and inequality constraints and connecting the model and its numerical solver to the objective function calculation as well as the optimizer (SQP). The default solver options in MoT can be kept and the desired variables to be plotted during the optimization need to be selected. Per default MoT plots the objective function, the parameter values and the dependent variables. After completion of the optimization MoT communicates the values of all model variables (in table), the value of the objective function and the generated plots can be exported to MS Excel. The modeller can access an automatically generated report containing the results from each step of the work-flow (and also the previously used work-flows) and statistics of the parameter estimation problem. In Phase IV the ratio of the values for the model discrimination criteria of scenarios 1 and 2 will be calculated and evaluated based on the resulting objective function value obtained in this step for each candidate scenario.

Step III.5: Statistical analysis of model predictions (Phase III)

The modeller can access an automatically generated report containing the results from each step of the work-flow (and also the previously used work-flows) and statistic of the parameter estimation problem.

5.5.4 Phase IV. Model evaluation/ validation

Evaluation

The performance of the candidate model is evaluated and compared based on the selected model discrimination measure in Phase III. In this case study the relative posterior probability of the candidate models $p^*\left(M_j|\boldsymbol{Y},\boldsymbol{\Sigma}\right)$ is calculated with the method proposed by Stewart et al. (1992, 1996, 1998) and applied for model discrimination. For this case study (case 2-b) the ratio of the posterior probabilities of two candidate models M_1 and M_2 with equal prior probability distribution is given by (see Appendix A7, Equation A7.7):

$$\frac{p^*(M_1|\boldsymbol{Y},\boldsymbol{\Sigma})}{p^*(M_2|\boldsymbol{Y},\boldsymbol{\Sigma})} = 2^{-\left(\frac{p_1}{2}-\frac{p_2}{2}\right)} \cdot \left(\frac{|\hat{v}_1|}{|\hat{v}_2|}\right)^{-\nu_e/2} \tag{5.825}$$

Here, $|\hat{v}_1|$ and $|\hat{v}_2|$ are the resulting objective function values from the parameter estimation performed during Phase III for scenario 1 and 2, respectively. The resulting value for the ratio of the posterior probabilities for model-scenario 1 and 2 is: 10^{228}. This means that scenario 1 is representing the experimental data by far better than scenario 2. The discrimination results between scenarios 1 and 2 are summarized in Table 5.44.

Table 5.44 *Terms of posterior probability of scenarios 1 and 2 (nomenclature see Appendix A7)*

	Scenario 1	***Scenario 2***
p_j	30	29
$\lvert\hat{v}_j\rvert$	1.63^{12}	9.04^{15}
v_e	163	163

The parameter values obtained during the parameter estimation in the previous phase for scenario 1 are given in Table 5.45.

Table 5.45 *Estimated parameter values for scenario 1*

Parameter:	***estimated value:***
kph	590.416
khp	502.565
kpa	0.89084
kap	234.118
$kpQ(1)$	9.091.923
$kQp(1)$	34.121
$kpQ(2)$	757.037
$kQp(2)$	0.29231
$kpQ(3)$	105.773
$kQp(3)$	32.125
$kpQ(4)$	1.438.748
$kQp(4)$	499.102
$kpQ(5)$	1.327.917
$kQp(5)$	0.861
$kpQ(6)$	1.365.568
$kQp(6)$	286.562
$kpQ(7)$	125.986
$kQp(7)$	0.29051
$kpQ(8)$	193.911
$kQp(8)$	0.69522
$kpQ(9)$	2.059.235
$kQp(9)$	16.568.218
$kpQ(10)$	1.397.443
$kQp(10)$	405.337
$kpQ(11)$	206.621
$kQp(11)$	0.41569
$kpQ(12)$	0.62555
$kQp(12)$	0.07646
$ke(1)$	0.30217
$ke(2)$	0.16006

Figure 5.49 shows a plot of the experimental data vs. the simulation results for the estimated parameter values of scenario 1. The dynamic profiles of the drug concentration in the kidney and the liver are shown in the figure.

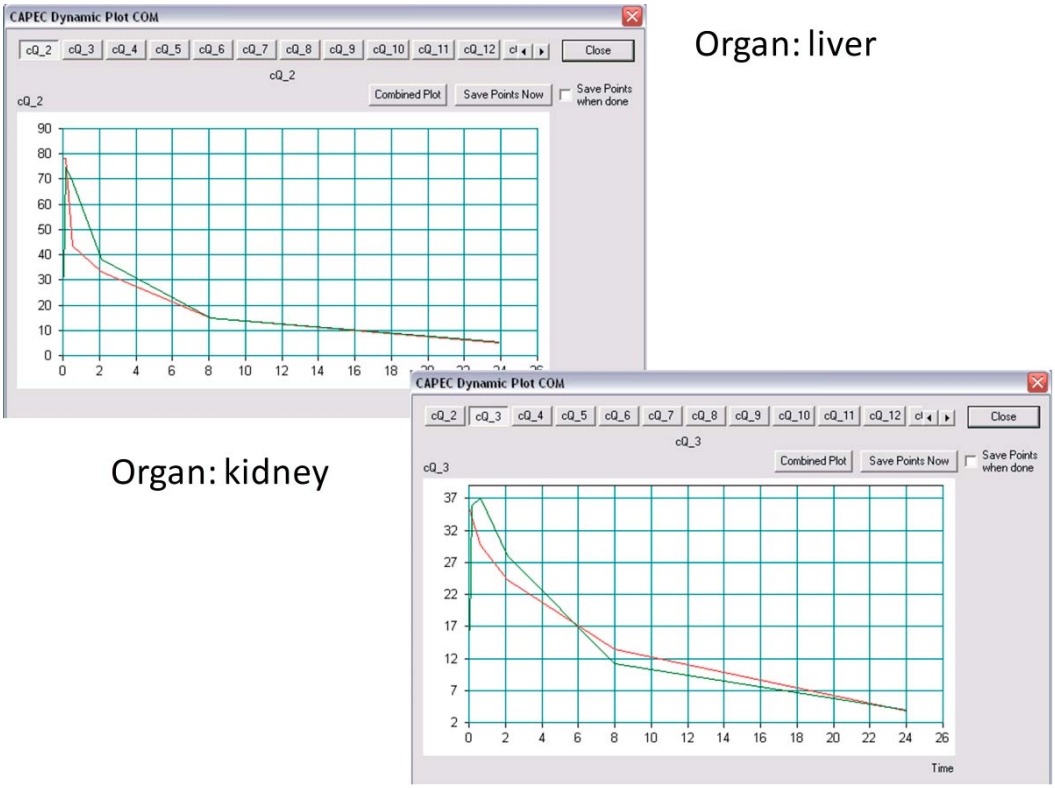


Figure 5.49 *Measured and simulated dynamic profiles of drug concentration in liver (top) and kidney (bottom) [µg/mL] after parameter estimation for scenario 1 generated by MoT.*

Potential for improving discrimination measure of scenarios based by simplification based on sensitivity and identifiability analysis

The results of the previous sensitivity analysis and identifiability analysis reveal that the potential for improving the posterior probability measure of the model candidates by model simplifications based on sensitivity analysis and identifiability analysis is not very high and of the same order of magnitude for both scenarios because for both scenarios only 2 parameters could be excluded.

5.5.5 2nd Model discrimination step for additional scenarios

In the previous section two candidate models with alternative mechanism for the drug transport between different compartments within the blood vessels have been discriminated. In a second model discrimination step alternative organ models are investigated. All candidate model-scenarios have the same blood vessel model because the optimal blood vessel model (scenario 1) has been identified in the first model discrimination step. The alternative organ models are derived based on scenario 1. Figure 5.50 shows the two additionally created scenarios together with scenario 1 (from previous discrimination step). Scenario 3 is a simplification of scenario 1 because it considers a 1-compartmental model for the rat organs. The compartment is assumed to be ideally mixed which means that the drug concentration is assumed to be the same in the entire organ. This assumption is justified by the available experimental data which provide an average concentration of the drug in the rat organs since the organs were homogenized before determination of the drug concentration. Scenario 3 considers a 2-compartment organ model like in scenario 1. For scenario 3 the equilibrium assumption for the transport of drug between interstitial fluid and tissue cells made by Mosat et al. (2011) (scenario 1) is dropped.

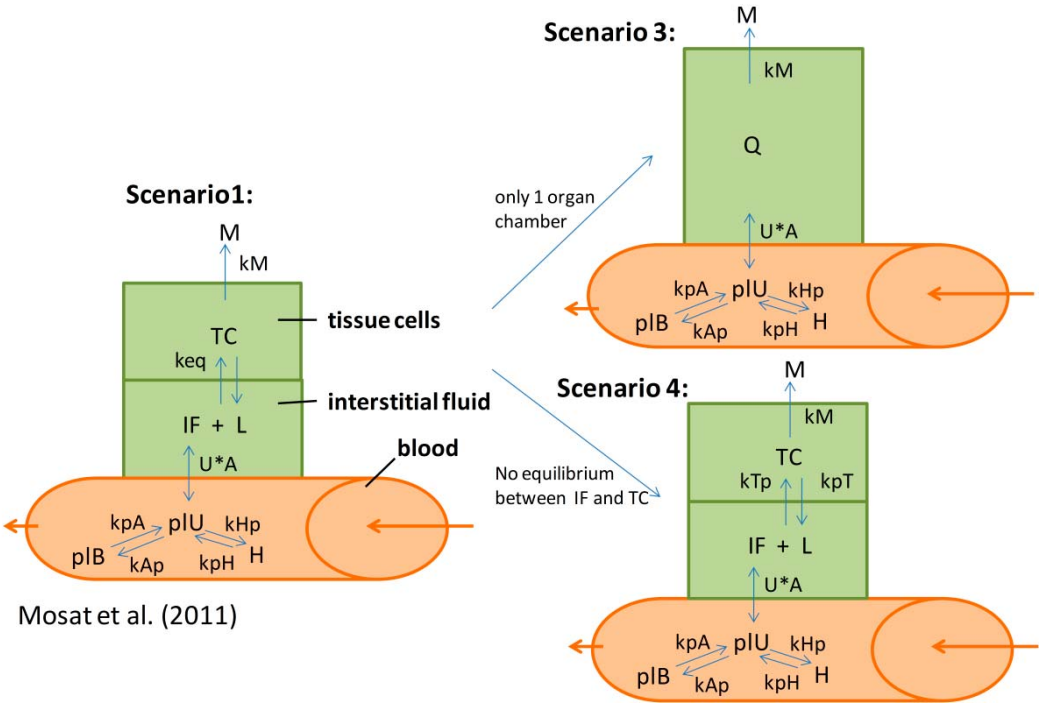


Figure 5.50 *Candidate model-scenarios with differing organ models*

The concepts for the new scenarios 3 and 4 are summarized in Tables 5.46 and 5.47, respectively.

Table 5.46 *Concept for scenario 3 .*

assumptions:	▪ Steady state blood flow ▪ Model blood flow by Hagen-Poiseuille ▪ 1-Compartment models for organs ▪ Organ compartment ideally mixed ▪ Simplified metabolism: irreversible 1st order reaction from drug to metabolite ▪ Metabolism occurs in liver and kidney (with different reaction rates) ▪ Mass transfer resistance between unbound plasma proteins and organ ▪ Consider only 12 organs (lung, heart, brain, bone, fat, muscle, skin, liver, spleen, guts, kidney, thymus) ▪ 3-Compartment model for blood vessel: hct, plasma bound, plasma unbound ▪ Volume of compartments for unbound drug in plasma, bound drug in plasma and hct are constant ▪ No saturation of blood cell binding (for the conditions present), modelled by 1st order reversible reactions
considered phenomena:	▪ Blood flow in main vessels (e.g. Hagen-Poiseuille) ▪ Multi-phase blood flow with dynamic transfer of drug between hct and plasma ▪ Dynamic binding of drug to plasma proteins in blood ▪ Mass transfer across cell membranes ▪ Metabolism of drug (reaction network)
scales:	***Low detail (macro scale):*** steady state blood flow network ***High detail (micro scale):*** dynamic distribution of drug considering a network with detailed organ models, multiphase blood-flow and metabolism

Table 5.47 *Concept for scenario 4 .*

assumptions:	▪ Steady state blood flow ▪ Model blood flow by Hagen-Poiseuille ▪ 2-Compartment models for organs: interstitial fluid and tissue cells ▪ Organ compartments ideally mixed ▪ Simplified metabolism: irreversible 1st order reaction from drug to metabolite ▪ Metabolism occurs in liver and kidney (with different reaction rates) ▪ Mass transfer of drug between IF and TC with ligands, 1st order reversible mass transfer of drug between IF and TC. ▪ Mass transfer resistance between unbound plasma proteins and interstitial fluid ▪ Consider only 12 organs (lung, heart, brain, bone, fat, muscle, skin, liver, spleen, guts, kidney, thymus) ▪ 3-Compartment model for blood vessel: hct, plasma bound, plasma unbound ▪ Volume of compartments for unbound drug in plasma, bound drug in plasma and hct are constant (from …, see Mosat et al., 2011) ▪ No saturation of blood cell binding (for the conditions present), modelled by 1st order reversible reactions
considered phenomena:	▪ Blood flow in main vessels (e.g. Hagen-Poiseuille) ▪ Multi-phase blood flow with dynamic transfer of drug between hct and plasma ▪ Dynamic binding of drug to plasma proteins in blood ▪ Mass transfer across cell membranes ▪ Mass transfer between interstitial fluid and cell tissue like a reversible 1st order reaction with a ligand ▪ Metabolism of drug (reaction network)
scales:	***Low detail (macro scale):*** steady state blood flow network ***High detail (micro scale):*** dynamic distribution of drug considering a network with detailed organ models, multiphase blood-flow and metabolism

The models for scenario 3 and scenario 4 need to be constructed following the model construction work-flow (Phase II) as previously done for scenarios 1 and 2. These steps are not shown here, but the model equations and numerical model analysis results for scenario 3 and scenario 4 are provided in Appendices A8 and A9, respectively. The 'Model identification/discrimination'-work-flow (Phase III) is applied for the discrimination between scenarios 1, 3 and 4. Also for this second model discrimination step the criteria proposed by Stewart et al. (1992, 1996, 1998) have been applied see (Appendix A7). In Phase IV, the performances of the three candidate model-scenarios are compared and evaluated based on the ratio of the posterior probabilities given by Equation 5.825. For scenario 4 and scenario 3 a posterior probability ratio of 10^{50} results. Consequently, the posterior probability of scenario 4 is remarkably higher than for scenario 3. This means that scenario 4 is representing the experimental by far better than scenario 3. The ratio of the posterior probabilities of scenario 4 and scenario 1 results to be 10^{-129}. So, in total scenario 1 is best supported by the available

experimental data. Table 5.48 summarizes the terms of Equation 5.825 for the three candidate models.

Table 5.48 *Terms for posterior probability of scenarios 1, 3 and 4 (nomenclature see Appendix A7)*

	Scenario 1	***Scenario 3***	***Scenario 4***
p_j	30	18	43
$\left\|\hat{v}_j\right\|$	1.63^{12}	6.48^{16}	3.22^{15}
v_e	163	163	163

Scenario 1 represents best the mechanisms for the drug transport between the different blood compartments (1st discrimination step) and for the organ model (2nd discrimination step) for the available experimental data. The model can be applied for

1. Simulation (monitor behaviour),
2. Optimization (e.g. optimize dosage curves)

following the corresponding model application work-flow described in Chapter 2.

5.5.6 Scale-up of rat model to human

The derived model for the distribution of CyA in the rat organism is to be scaled-up to a human (see Figure 5.51).

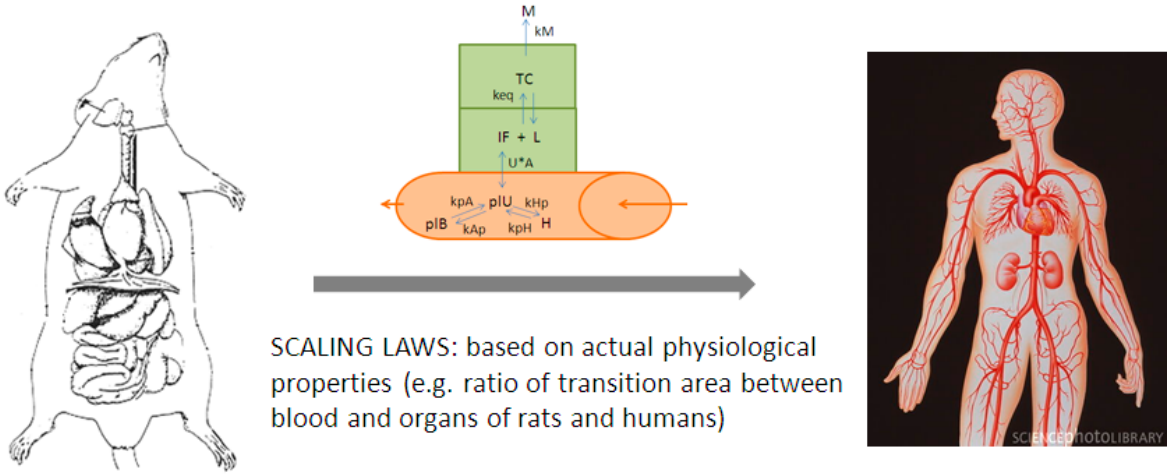


Figure 5.51 *Scale-up of rat model to human being*

The scale-up is motivated by the fact that only limited experimental data is accessible for humans because it is not possible to measure the dynamic drug concentration profiles in the organs. The dynamic blood concentration profiles can be accessed experimentally. The scale-up

of the rat model (scenario 1) is performed based on physiologically based scaling laws. The transfer coefficients between blood vessel (plU) and interstitial fluid are, for example, scaled up based on the ratio of the phase transition areas between these compartments in rats and humans. Furthermore, the organ volumes, the blood volume in each organ, the total blood volume in the body, the blood flow rates and the total mass of the organism are adjusted to a human weighing 73 kg. More detailed information on the applied scaling laws is given in Hall et al. (2011) and Kawai et al. (1998).

Figure 5.52 shows the available experimental data for blood and plasma concentrations available for humans in comparison with the simulation results obtained for the scaled-up human model without re-fitting any parameters.

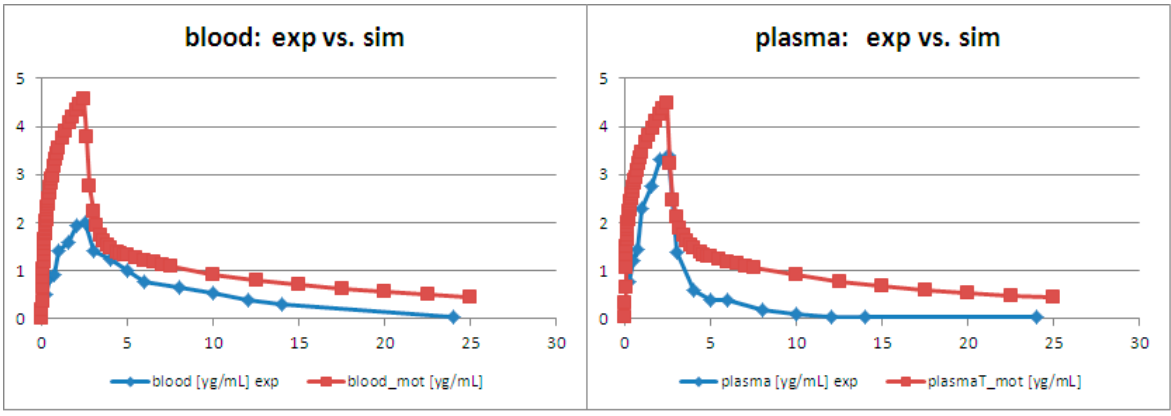


Figure 5.52 *Scaled-up human model: Experimental data vs. simulation results for measured dynamic drug concentration in blood and plasma.*

In order to get an idea which parameters of the scaled-up human model can be re-fitted with the limited human data available sensitivity and identifiability analysis are very useful tools. Figure 5.53 gives the obtained overall parameter significance ranking from a local differential sensitivity analysis performed at the conditions and for the measured variables of all available experimental data-points. The perturbation of the parameters has been set to 0.1%.

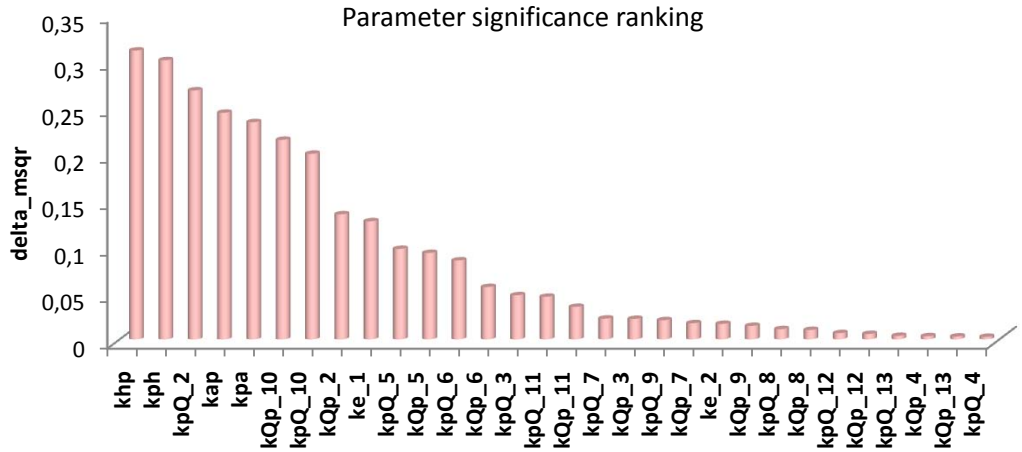


Figure 5.53 *Overall parameter significance ranking for all 30 model parameters of scaled-up human model.*

The top 18 parameters were selected for further analysis. The results of the performed identifiability analysis are summarized in Table 5.49. The threshold assigned to the collinearity index is 5. Table 5.49 for each subset size (1st column) shows the number of possible parameter subsets, the number of identifiable subsets (highlighted column), as well as the minimum and maximum collinearity indexes (γ).

Table 5.49 *Results of identifiability analysis for scaled-up human model and top 18 most sensitive model parameters (resulting from previous sensitivity analysis step): Statistics of identifiable parameter subsets*

subset size	*no. of subsets*	*no. of identi-fiable subsets*	*min. γ*	*max. γ*
2	153	140	1.00	94.70
3	816	538	1.33	350.31
4	3060	1000	1.64	720.25
5	8568	832	2.57	1653.59
6	18564	236	4.53	3799.77

The maximal parameter subset size that is identifiable (collinearity index lower than the threshold of 5) by the limited available human data is 6. The 6-parameter subset with the lowest collinearity index is selected for fine-tuning applying the limited human data.

5.5.7 Conclusions for case study

The pharmacokinetic case study has highlighted the application of computer-aided work-flows in MoT for Phases I, II, III and IV. In Phase II both, the multi-scale model construction work-flow (for the construction of the overall multi-scale scenario) and the single-scale model construction work-flow (for the construction of the models for the different scales within the multi-scale scenarios) has been demonstrated. In Phase III a model discrimination measure based on Stewart et al. (1992, 1996, 1998) has been applied, which in addition to the objective function value obtained in the parameter estimation step also includes a penalty on the number of model parameters. Different candidate models with alternative model scenarios for the blood vessels/blood flow as well as the rat organs have been evaluated and compared in Phase IV based on the selected model discrimination measure (Phase III). The final model scenario considers a 2-compartment organ model consisting of interstitial fluid (IF) and cell tissue (CT). The transport between these compartments is assumed to be at equilibrium whereas the transport between the blood vessel and the interstitial fluid is dependent on the concentration difference between these compartments. For the blood vessel segments a 3-compartment model is applied consisting of a compartment for the red blood cells, the plasma bound and the plasma unbound. The transport of drug between these compartments is modelled as reversible 1^{st} order reactions. The developed final model has been scaled up for a human being applying advanced scaling laws. Using the sensitivity and identifiability analysis tools in MoT has allowed to identify which parameter subsets of the scaled-up human model can be re-fitted applying the limited experimental data accessible for humans (dynamic organ concentration profiles not accessible). The systematic modelling strategy based on computer-aided work-flows for model documentation, construction, discrimination and evaluation together with the pharmacokinetic models available in the MoT model library increases the efficiency of the modeller in the derivation of pharmacokinetic models appropriate with respect to a modelling goal and accessible experimental data.

Chapter 6. Discussion

6.1 Summary of main contribution of PhD thesis

The main contributions of this Ph.D. thesis are summarized by:

- **Development of an overall modelling methodology based on in-depth work-flows and data-flows for the different generic modelling tasks required for model development, analysis, identification, discrimination, documentation and application for simulation and optimization** *(see Chapter 2)*
    3. Identification of modelling tasks and their interconnection (based on literature, existing methodologies for the modelling process, case studies).
    4. Development of in-depth work-flows and data-flows for the modelling tasks including required methods (based on literature, existing work-flows and methods).
- **Development of a computer-aided modelling framework that is structured based on the generic modelling methodology by elaborating how the computer can support the modeller and making the identified work-flows and data-flows computer-aided** *(see Chapter 3)*
    4. Identification of required tools and features for each step in order to provide maximum support to the modeller.
    5. Provision of expertise and insight on theoretical background and application of the different methods and tools.
    6. Analysis of opportunities for automation of steps without restriction of flexibility.
- **Development of a software architecture and implementation of the computer-aided modelling framework into user-friendly software** *(see Chapter 4)*
    - Modification and extension of the existing ICAS-MoT modelling tool (Russel & Gani, 2000, Sales-Cruz & Gani, 2003, Sales-Cruz, 2006) according to developed computer-aided modelling framework and software architecture.
    - Methods and tools needed for the computer-aided modelling framework (see Chapter 3) have been implemented and are summarized in Table 6.1 (and Appendix A1).
- **Solution of case studies from very different areas in chemical and biochemical engineering** *(see Chapter 5)*
    - Validation and demonstration of proposed work-flow based modelling methodology.

- Validation and demonstration of developed computer-aided modelling framework/ modelling software.
- Figure 6.1 gives an overview of the main case studies which have been solved in ICAS-MoT in the scope of this project.

- **Demonstration how templates for specific application areas are compatible with the generic work-flows and how these templates can be integrated in the computer-aided modelling framework in order to provide a maximum of domain knowledge and support but at the same time keep up the flexibility of the generic framework** (fragrance spraying example)

The benefits of the developed computer-aided modelling framework can be summarized as follows:

- Systematisation of the modelling process; provision of structure, guidance and support;
- Increase of efficiency of process of model development and application with respect to time and resources (combination of required methods and tools, storage of knowledge, documentation, re-use of models and knowledge, automation);
- Incorporation of state-of-the-art methods;
- Increase model quality and reliability.

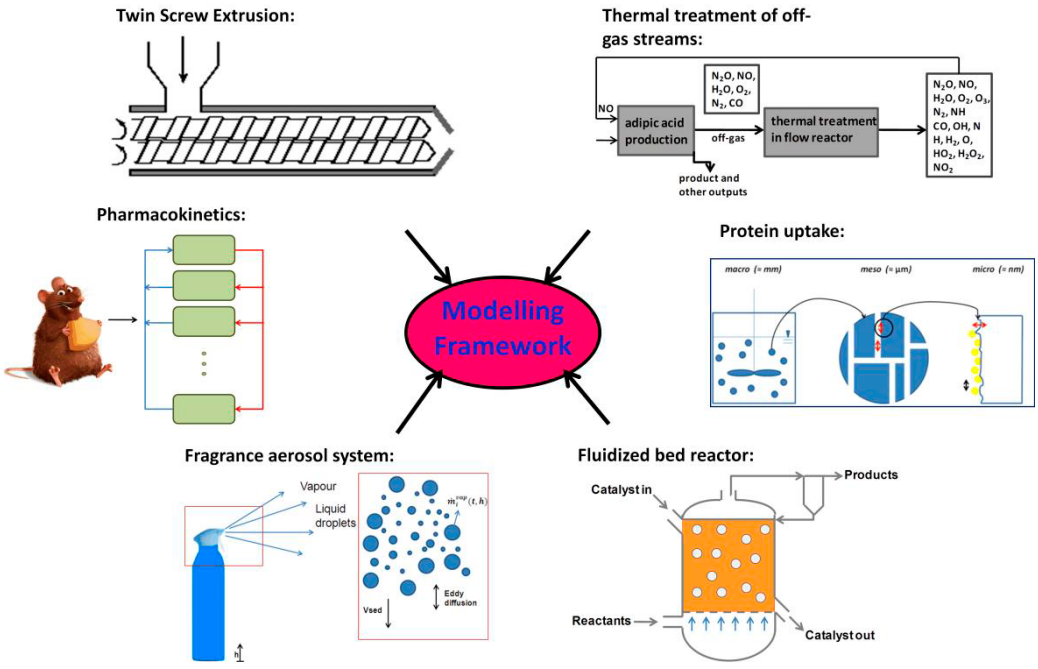


Figure 6.1 *Overview of case studies for development and validation of computer-aided modelling framework.*

To evaluate the actual implementation of the computer-aided modelling framework in the ICAS-MoT modelling tool some key issues need to be considered. Table 6.1 gives a list of these issues in terms of key features of a modelling tool as identified in Section 1.3 (Table 1.4). For each key feature, the ones actually available in MoT are also listed. A large portion of the MoT-features were already implemented prior to the start of this Ph.D.-project. In order to indicate which features have been entirely implemented within the scope of this project, they are underlined with a solid line (and/or red letters). The features which have been partly implemented or extended during the Ph.D.-project are shown in italic (and/or yellow letters). The features which were already available at the start of the project are not underlined (and/or black letters). Appendix A1 gives a more detailed overview on the features implemented in MoT during this Ph.D.-project.

Table 6.1 *Key areas for a modelling tool to provide support (identified during literature review in Section 1.3, Table 1.4) and corresponding features available in MoT*

Key features of modelling tool	***MoT***
1 Structured development of models, automation of part of the modelling process	Computer-aided, partly automated work-flows for the different modelling tasks*
2 Model documentation	▪ structured documentation interface ▪ automated documentation and report generation of all conducted work-flow steps ▪ in-line comments with model equations*
3 Model re-use	▪ *model documentation features** ▪ save model files ▪ *model library*
4 Model libraries	*library of models for different applications*
5 Model decomposition	▪ linking and data-flow schemes ▪ incidence matrix analysis ▪ eigenvalue report (identifies differing time-scales)
6 Model aggregation	▪ linking and data-flow interface ▪ additional equations can be copied to MoT equation editor ▪ *call of sub-models*
7 Support for equation generation	▪ link to ICAS-ModDev ▪ equation derivation is considered in algorithmic work-flow for model construction ▪ *model libraries and aggregation (re-use)*
8 Simple implementation of model equations	▪ provision of model equations in simple txt-syntax and translation (including partly automated numerical model analysis) by RPN ▪ *automated PDE discretization* ▪ computer-aided model construction work-flows

Table 6.1 *CON'T (Key areas for a modelling tool to provide support (identified during literature review in Chapter 1.3, Table 1.4) and corresponding features available in MoT)*

Key features of modelling tool	***MoT***
9 Numerical model analysis	▪ variable classification (partly automated) ▪ degree of freedom analysis (in-line with variable classification) ▪ singularity check (in-line with variable classification) ▪ incidence matrix generation and analysis ▪ optimized equation ordering ▪ derivation of solution strategy and connection of required solvers ▪ eigenvalue report ▪ computer-aided model construction work-flows
10 Model verification/ debugging	▪ numerical model analysis ▪ singularity check ▪ debugger (equation by equation) ▪ equation generation, model libraries, *model aggregation*
11 Systematic model reduction/ simplification	▪ eigenvalue report ▪ derivation and output of (exact) Jacobian matrix
12 Model identification/ validation	▪ computer-aided model identification work-flow ▪ *sensitivity analysis* ▪ identifiability analysis ▪ optimizer ▪ *statistics* ▪ uncertainty analysis
13 Simulation	▪ computer-aided simulation work-flow ▪ automated connection of appropriate numerical solvers ▪ *numerical solvers* ▪ graphical interface to adjust solver options ▪ graphical interface to select variables to plot ▪ eigenvalue report ▪ *automated report generation* ▪ output of variable value tables and plots ▪ run MoT from MS-Excel and export results to MS-Excel
14 Optimization	▪ computer-aided optimization work-flow ▪ automated connection optimizer ▪ optimizer ▪ graphical interface to adjust solver options ▪ *automated report generation* ▪ output of variable value tables and plots

Table 6.1 *CON'T (Key areas for a modelling tool to provide support (identified during literature review in Chapter 1.3, Table 1.4) and corresponding features available in MoT)*

Key features of modelling tool	***MoT***
15 Support for multi-scale modelling	▪ computer-aided multi-scale model construction work-flow ▪ data-flow and linking schemes and corresponding interface ▪ eigenvalue report ▪ iterative determination of required degree of detail/complexity with respect to modelling goal and validation potential ▪ connection to ICAS process simulation ▪ ICAS-ProPred (provided with ICAS software package but no direct link)
16 Domain knowledge/support	▪ *Model library* (in connection with structured documentation interface and automated report generation) ▪ Connection to thermodynamic data-base of ICAS ▪ Export of models to process simulator in ICAS ▪ Provision of MoT together with ICAS software (-> easy access to other ICAS tool not directly integrated with MoT, e.g. property prediction, solvent design)

* - features implemented during Ph.D.-project

* - features already available in MoT at the start of Ph.D-project

** - features partly implemented/extended during Ph.D.-project*

6.2 Open challenges

6.2.1 Generic modelling methodology and computer-aided modelling framework

In-depth work-flows for sub-tasks within the modelling process, for example, design of experiments for model discrimination or parameter estimation, need to be included in the generic modelling methodology and computer-aided modelling framework. Furthermore, a computer-aided in-depth work-flow for model evaluation and validation needs to be developed.

6.2.2 Modelling tool (ICAS-MoT)

There is a vast potential for extending and improving ICAS-MoT. Here, the main issues and needs are summarized:

- Implementation of missing features and automation identified during development of computer-aided modelling framework (Chapter 3):
    - Support more linking scheme types and features;
    - Solution of sub-models at different time-scales;
    - Multiple-scenario manager.
- More alternative methods and choices, e.g. for:
    - Objective function selection/model discrimination measure selection (for example Bayesian parameter estimation and/or model discrimination);
    - Sensitivity analysis methods;
    - Implementation of alternative optimizers to SQP solver (e.g. MINLP-solver, global optimizer);
    - Extension of available PDE discretization methods, especially for automated discretization in 2 or 3 dimensions besides time.
- Extension and improved structure of library including a large variety of different models for properties, phenomena, unit operation and other systems and processes;
- Automated derivation of steady state or linearized models;
- Option to provide units for each variable during numerical model analysis and corresponding automated unit conversion and checks;
- Implementation of developed fragrance aerosol template and similar templates for different systems (-> template library).

Appendix

A1. List of improvements and changes in MoT

A. Interface

- work-flow interface (work-flows for modelling objective and system information, single- and multi-scale model construction, model identification, optimization and simulation)
- automation of work-flow steps when possible (e.g. set-up of sensitivity analysis, experimental data-interface, change from optimization to simulation problem)
- explanations of different work-flow steps and tools
- link to ICAS-ModDev

B. Model documentation

- structured documentation interface
- automated report generation from all work-flow steps performed (new reports for model construction, model identification, optimization, eigenvalue analysis, sensitivity analysis, overall report) (generation of a statistic report of model solution was already available)

C. Partial differential equations (PDEs)

- Method of lines: modeller can enter desired number of discretization steps
- Discretization of coupled PDEs
- Orthogonal collocation (symmetric problems)

D. Model Analysis:

- Eigenvalue report generation for ODE and AE systems
- Local differential sensitivity analysis:
    - Parameter significance ranking
    - Multiple runs feature (change of variable/parameter values between runs) and presentation of all results in tables, overall parameter significance ranking from all runs)

  - For dynamic problems: calculation of sensitivities at multiple output times and presentation of all results in tables, overall parameter significance ranking from all output times (and runs); different output times possible for different response variables
  - Plot of sensitivity functions: normalized sensitivity vs. run and/or time (for selected parameters and response variables)
  - Automated set-up of sensitivity analysis for optimization and model identification (based on experimental data)
- Global sensitivity analysis (Morris screening)
- Identifiability analysis
- Calculation of confidence intervals, correlation matrix and variance matrix
- Monte Carlo uncertainty analysis

E. Solver and Solver options:

- Gauss solver for linear algebraic equation systems based on derivation of exact Jacobian from entered model equations
- Solution of dynamic systems

  a) Solver options: relation of number of output timesteps, final integration time and scaling factor, modeller can provide final value of independent variable

  b) Extension of multiple run feature: change model parameters and/or solver specifications between runs

F. Experimental data-interface and parameter estimation:

- possible to select between different objective functions for parameter estimation from a displayed list of options; applied objective function is provided in reports
- modeller can include weight factors for each data point in the objective function
- information from experimental data-interface is automatically passed over to sensitivity analysis and parameter estimation
- dynamic parameter estimation for all possible values of independent variable

G. General handling of MoT

- Mot-syntax: Handle vector and matrix elements j+integer, j-integer
- 4 blank lines after translation will not be created anymore
- save buttons for: perturbation set-up, design variables, constraint variables and variable chart trace
- new MoT example files in CAPEC folder

- thermo-functions also work if numbering of vector elements for vectors that are arguments in thermo-functions does not start from 0 but from any value

E. Linking of models/ specific multi-scale features

- data-flow and linking scheme interface with adjustment of variable sized between linked models and with respect to number of calls
- multiple calls of a sub-model from a main-model
- call sub-models with ODEs
- matrices can be arguments in sub-model calls
- provide solver options for sub-models
- sub-function to calculate the x%-statistic diameter of a droplet/particle size distribution

A2. Experimental data for case study 1 (thermal treatment of off-gas stream of adipic acid production)

feed [ppm]:	dataset 1:		dataset 2:		dataset 3:		dataset 4:		dataset 5:	
	H2O=100, N2O=227, Ar=balance		H2O=10000, N2O=201, Ar=balance		H2O=150000, N2O=227, Ar=balance		H2O=10, N2O=196, O2=balance		H2O=100, N2O=194, CO2=balance	
	T[K]	N2O [PPM]	T[K]	N2O [PPM]	T[K]	N2O [PPM]	T[K]	N2O [PPM]	T[K]	N2O [PPM]
	1008	226	1091	199	1013	199	1025	195	1026	190
	1034	224	1115	197	1041	196	1047	196	1050	190
	1058	225	1140	195	1060	196	1072	193	1072	188
	1084	224	1165	193	1086	196	1095	191	1098	186
	1108	223	1193	189	1110	195	1121	189	1121	183
	1134	221	1217	182	1136	191	1145	187	1146	180
	1159	218	1242	174	1160	189	1171	184	1170	176
	1186	213	1265	162	1188	182	1198	180	1199	169
	1210	205	1288	144	1213	173	1222	173	1223	159
	1235	194	1312	123	1236	160	1245	163	1245	146
	1259	178	1337	101	1259	143	1269	150	1271	126
	1284	156	1353	74	1284	121	1294	133	1294	102
	1302	132	1381	49	1306	96	1319	111	1318	76
	1320	106			1324	72	1344	85	1342	48
	1347	73					1367	58	1366	24
							1389	33	1390	8

* The residence times at 1200 K (constant molar rate) are 188 ms (set1), 157 ms (set 2), 160 ms (set 3), 158 ms (set 4) and 158 ms (set 5), data taken from Glarborg et al. (1994)

A3. Model equations case study 1 (thermal treatment of off-gas stream of adipic acid production)

(A3.1) $Vp = \frac{R \cdot T}{P} \cdot \sum_{j=1}^{NC} F[j]$

(A3.2-A3.16) $H^0[j] = a1[j] + a2[j] \cdot \frac{T}{2} + a3[j] \cdot \frac{T^2}{3} + a4[j] \cdot \frac{T^3}{4} + a5[j] \cdot \frac{T^4}{5} + \frac{a6[j]}{T}\ ,\ j = 1 \dots NC - 3$

(A3.17-A3.31) $S^0[j] = a1[j] \cdot \ln(T) + a2[j] \cdot T + a3[j] \cdot \frac{T^2}{2} + a4[j] \cdot \frac{T^3}{3} + a5[j] \cdot \frac{T^4}{4} + a7[j], j = 1 \dots NC - 3$

(A2.32-A3.75) $HR[k] = \sum_{j=1}^{NC} \nu_{kj} \cdot H^0[j]\ ,\ k = 1 \dots NR$

(A3.76-A3.119) $SR[k] = \sum_{j=1}^{NC} \nu_{kj} \cdot S^0[j]\ ,\ k = 1 \dots NR$

(A3.120-A3.163) $K[k] = \exp\{SR[k] - HR[k]\} \cdot \left(\frac{1}{RT}\right)^{\sum_{j=1}^{NC} \nu_{kj}}, k = 1 \dots NR$

(A3.164- A3.203) $kf[k] = A[k] \cdot T^{\beta[k]} \cdot e^{\frac{-E[k]}{RT}},\ k = 1 \dots NR, k \neq 7, 8, 9, 23$

(A3.204) $F_{M2} = F[1] + F[3] + F[4] + F[5] + F[7] + F[8] + F[9] + F[11] + F[12] + F[13] + F[14] + F[15] + F[16] + F[17] + F[18]$

(A3.205) $F_{M6} = F[1] + F[2] + F[3] + F[4] + 5F[5] + F[6] + F[7] + F[8] + F[9] + F[10] + F[11] + F[12] + F[13] + F[14] + F[15] + F[16] + F[17] + F[18]$

(A3.206) $F_{M7} = F[1] + 2F[2] + F[3] + 0{,}78F[4] + F[5] + 11F[6] + F[7] + F[8] + F[9] + F[11] + F[12] + F[13] + F[14] + F[15] + F[17] + F[18]$

(A3.207) $F_{M10} = F[1] + F[2] + F[3] + 1{,}5F[4] + F[5] + 10F[6] + F[7] + F[8] + F[9] + 1{,}5F[10] + F[11] + F[12] + F[13] + F[14] + F[15] + F[16] + F[17] + F[18]$

(A3.208) $F_{M12} = F[1] + 2{,}5F[2] + F[3] + F[4] + F[5] + 12F[6] + F[7] + F[8] + F[9] + F[10] + F[11] + F[12] + F[13] + F[14] + F[15] + F[16] + 1{,}9F[17] + 3F[18]$

(A3.209) $F_{M23} = F[1] + 2{,}5F[2] + F[3] + F[4] + F[5] + 12F[6] + F[7] + F[8] + F[9] + F[10] + F[11] + F[12] + F[13] + F[14] + F[15] + 0{,}64F[16] + F[17] + F[18]$

(A3.210) $F_M = \sum_{j=1}^{NC} F[j]$

(A3.211) $kinf7 = A[7] \cdot T^{\beta[7]} \cdot e^{\frac{-E[7]}{RT}}$

(A3.212) $kinf8 = A[8] \cdot T^{\beta[8]} \cdot e^{\frac{-E[8]}{RT}}$

(A3.213) $kinf9 = A[9] \cdot T^{\beta[9]} \cdot e^{\frac{-E[9]}{RT}}$

(A3.214) $kinf23 = A[23] \cdot T^{\beta[23]} \cdot e^{\frac{-E[23]}{RT}}$

(A3.215) $klow7 = Alow7 \cdot T^{\beta low7} \cdot e^{\frac{-Elow7}{RT}}$

(A3.216) $klow8 = Alow8 \cdot T^{\beta low8} \cdot e^{\frac{-Elow8}{RT}}$

(A3.217) $klow9 = Alow9 \cdot T^{\beta low9} \cdot e^{\frac{-Elow9}{RT}}$

(A3.218) $klow23 = Alow23 \cdot T^{\beta low23} \cdot e^{\frac{-Elow23}{RT}}$

(A3.219) $X7 = \frac{FM7 \cdot klow7}{kinf7 \cdot Vp}$

(A3.220) $X8 = \frac{FM7 \cdot klow8}{kinf8 \cdot Vp}$

(A3.221) $X9 = \frac{FM9 \cdot klow9}{kinf9 \cdot Vp}$

(A3.222) $X23 = \frac{FM23 \cdot klow23}{kinf23 \cdot Vp}$

(A3.223) $Fcent7 = (1 - aa7) \cdot e^{\frac{-T}{T7^{***}}} + aa7 \cdot e^{\frac{-T}{T7^*}}$

(A3.224) $Fcent8 = (1 - aa8) \cdot e^{\frac{-T}{T8^{***}}} + aa7 \cdot e^{\frac{-T}{T8^{*}}}$

(A3.225) $Fcent9 = (1 - aa9) \cdot e^{\frac{-T}{T9^{***}}} + aa7 \cdot e^{\frac{-T}{T9^{*}}}$

(A3.226) $Fcent23 = (1 - aa23) \cdot e^{\frac{-T}{T23^{***}}} + aa7 \cdot e^{\frac{-T}{T23^{*}}}$

(A3.227) $c7 = -0{,}4 - 0{,}67 \cdot \log(Fcent7)$

(A3.228) $c8 = -0{,}4 - 0{,}67 \cdot \log(Fcent8)$

(A3.229) $c9 = -0{,}4 - 0{,}67 \cdot \log(Fcent9)$

(A3.230) $c23 = -0{,}4 - 0{,}67 \cdot \log(Fcent23)$

(A3.231) $N7 = 0{,}75 - 1{,}27 \cdot \log(Fcent7)$

(A3.232) $N8 = 0{,}75 - 1{,}27 \cdot \log(Fcent8)$

(A3.233) $N9 = 0{,}75 - 1{,}27 \cdot \log(Fcent9)$

(A3.234) $N23 = 0{,}75 - 1{,}27 \cdot \log(Fcent23)$

(A3.235) $F7 = 10^{\frac{\log(Fcent7)}{1+\left(\frac{\log(X7)+c7}{N7-0{,}14\cdot(\log(X7)+c7)}\right)^2}}$

(A3.236) $F8 = 10^{\frac{\log(Fcent8)}{1+\left(\frac{\log(X8)+c8}{N8-0{,}14\cdot(\log(X8)+c8)}\right)^2}}$

(A3.237) $F9 = 10^{\frac{\log(Fcent9)}{1+\left(\frac{\log(X9)+c9}{N9-0{,}14\cdot(\log(X9)+c9)}\right)^2}}$

(A3.238) $F23 = 10^{\frac{\log(Fcent23)}{1+\left(\frac{\log(X23)+c23}{N23-0{,}14\cdot(\log(X23)+c23)}\right)^2}}$

(A3.239) $kf[7] = \frac{kinf7 \cdot X7 \cdot F7}{1+X7}$

(A3.240) $kf[8] = \frac{kinf8 \cdot X8 \cdot F8}{1+X8}$

(A3.241) $kf[9] = \frac{kinf9 \cdot X9 \cdot F9}{1+X9}$

(A3.242) $kf[23] = \frac{kinf23 \cdot X23 \cdot F23}{1+X23}$

(A3.243- A3.286) $kb[k] = \frac{kf[k]}{K[k]}\,, \quad k = 1 \dots NR$

(A3.287- A3.330) $r[k] = kf[k] \cdot \prod_{j=1}^{NC} \left(\frac{F_{reactant,j}}{Vp}\right)^{|\nu_{kj}|} - kb[k] \cdot \prod_{j=1}^{NC} \left(\frac{F_{product,j}}{Vp}\right)^{\nu_{kj}}, \quad k = 1 \dots NR$

(A3.331- A3.345) $dF[j] = \sum_{k=1}^{NR} \nu_{kj} \cdot r[k]\,,\ j = 1 \dots NC\text{-}3$

Material balances are needed for all 15 non-inert compounds j (equations 331-345*)*. The remaining equations are needed to provide expressions for the reaction rate $r[k]$ (equations 287-330) of reaction k which are needed for the RHS of the material balances. The rate constants for the forward reactions $kf[k]$ are calculated by the Arrhenius equation (equations 164-203). The rate constants of the backward reactions $kb[k]$ are resulting as the ratio of the equilibrium constants and the rate constants of the forward reactions (equations 343-286). The equilibrium constants $K[k]$ of the reactions are calculated from the standard reaction enthalpies and entropies (equations 120-163). The compound standard enthalpies $H[j]$ and entropies $S[j]$ at the system temperature are determined with the Nasa-Polynomials (equations 2-31). They are obtained in dimensionless from. The reaction enthalpies $HR[k]$ and entropies $SR[k]$ then result from stoichiometry (equations 32-119). The rate constant for the forward reactions kf is calculated in equations 164-203 and 239-242. For

certain reactions the third body enhancement needs to be considered. The concentrations of the third body for these reactions are calculated in equations 204-210. Here, F_{M2} is the concentration of the third body in reaction 2, etc. For reactions 8, 9, 37, 40, 44 all enhancement factors are 1 and the concentration of the third body is simply the total concentration in the reactor F_M. Tor the third body reactions 7, 8, 9 and 23 the pressure dependence of the rate constant needs to be considered according to the Troe's equations. The needed parameters for the Troe equation are given by equations 211-238.

A4. Sensitivity functions for all possible parameter pairs case study 1 (thermal treatment of off-gas stream of adipic acid production)

For this case study the sensitivity functions for a parameter cannot be plotted in the same diagramm because the experimental data has been aquired at different feed conditions for each dataset. So for each parameter pair 5 diagrams have been plotted for the five datasets. These five diagrams are presented in one row for the same parameter pair. If the curves of the sensitivity functions are collinear in all 5 diagrams the two parameters are correlated.

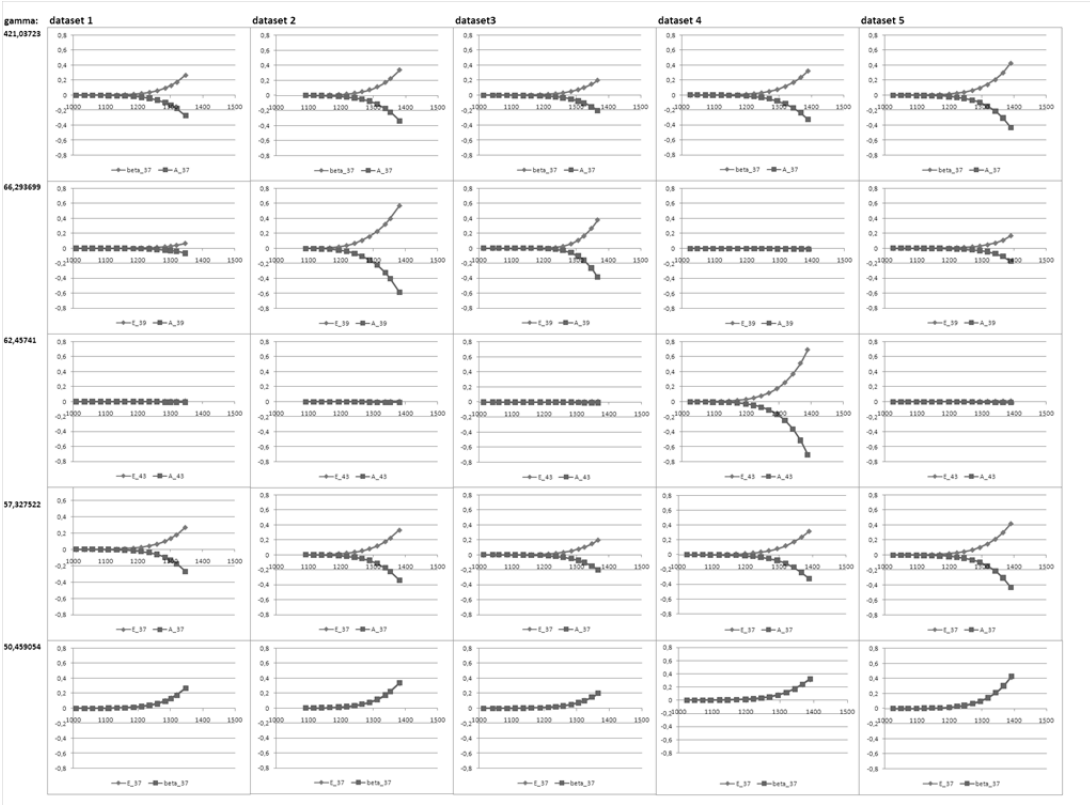
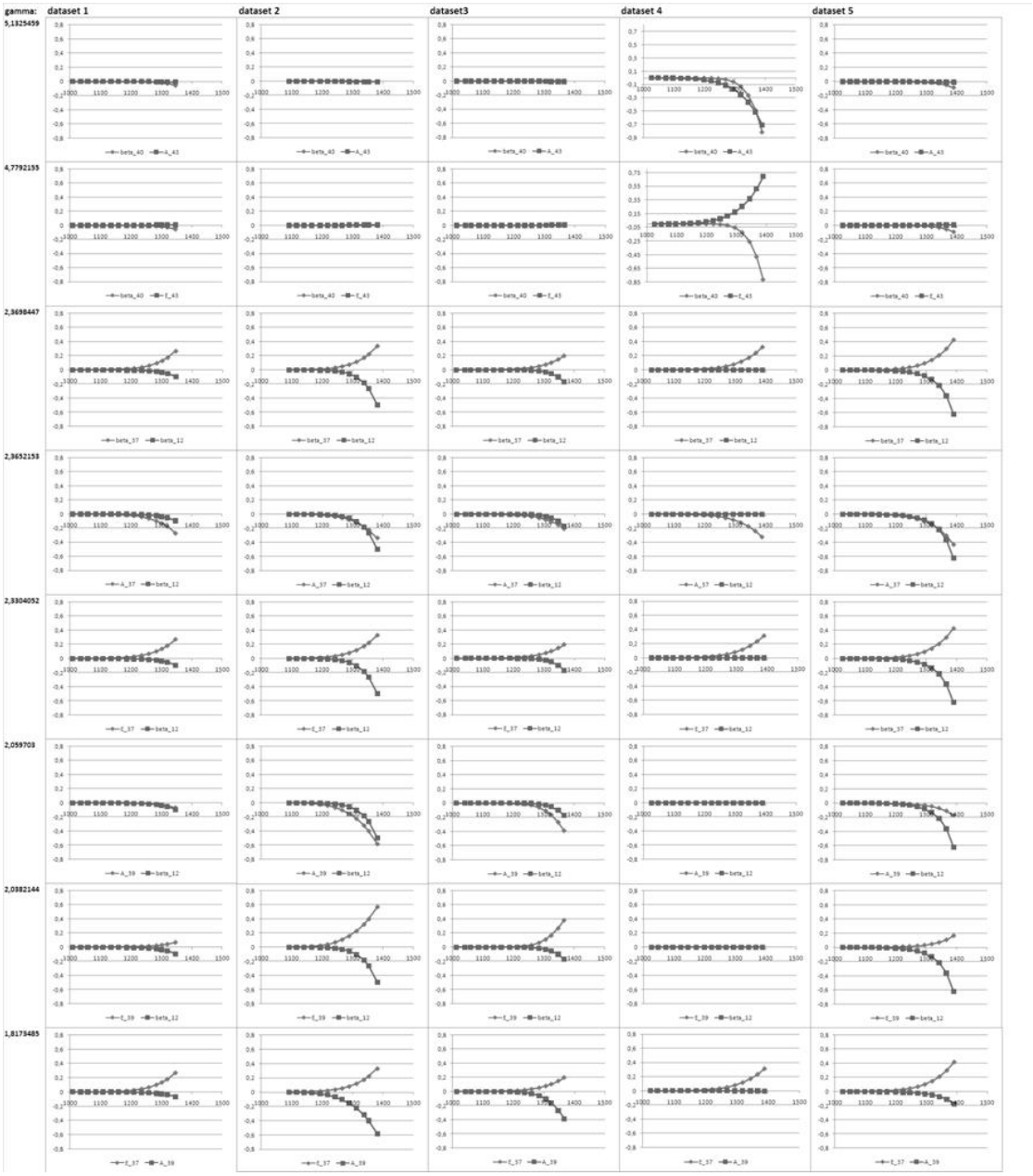
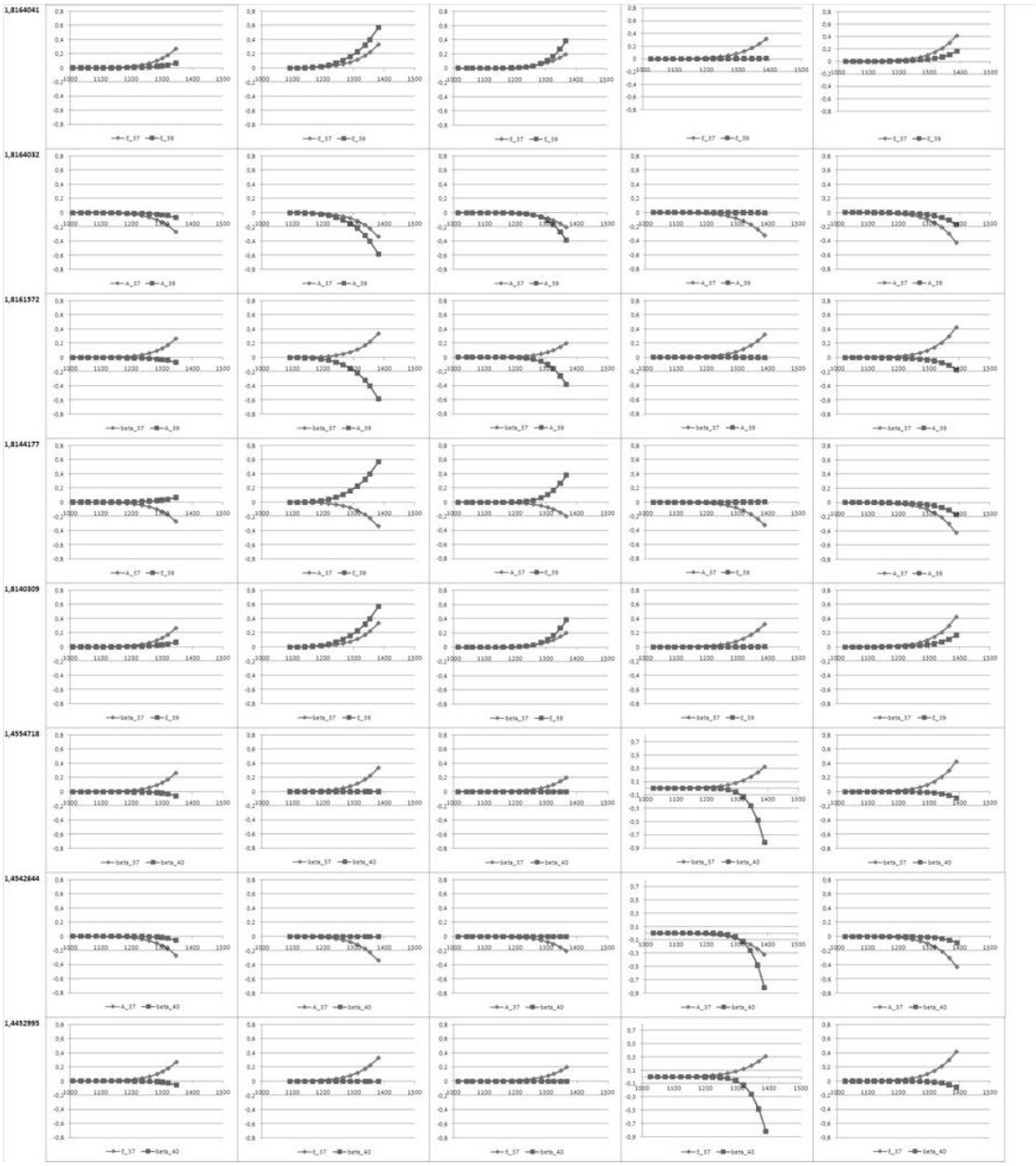
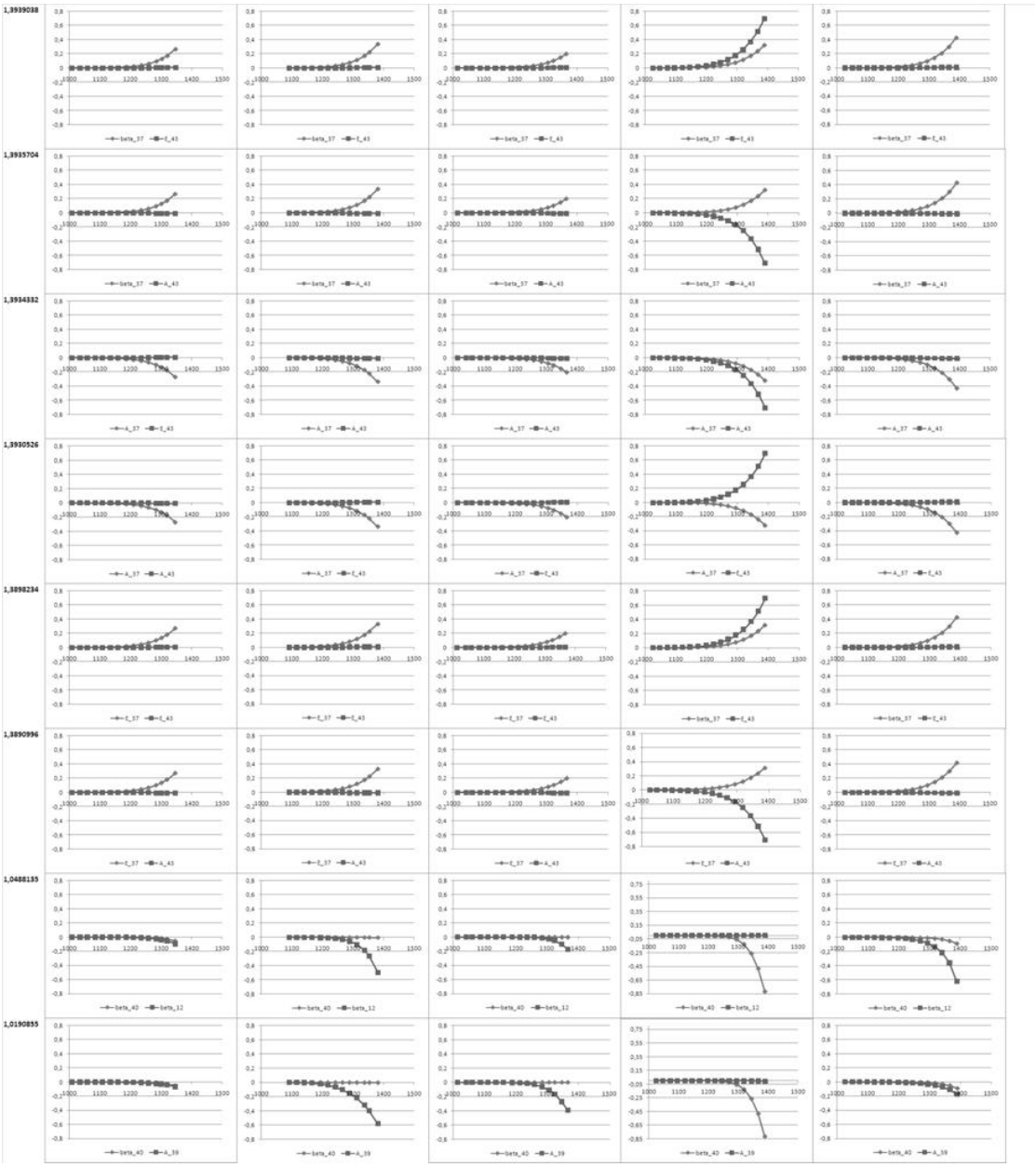


Figure A4.1 *Comparison of sensitivity functions of parameter pairs (here for all parameter pairs that result to be not-identifiable in the later quantitative analysis, see Chapter 5.1, Phase II, Step 3).*


gamma:
dataset 1
dataset 2
dataset3
dataset 4
dataset 5
5,1325459
4,7792155
2,3698447
2,3652153
2,3304052
2,059703
2,0382144
1,8173485


1,8164041
1,8164032
1,8161572
1,8144177
1,8140309
1,4554718
1,4542644
1,4452995


1,3939038
1,3935704
1,3934832
1,3930526
1,3898234
1,3890996
1,0488135
1,0190855

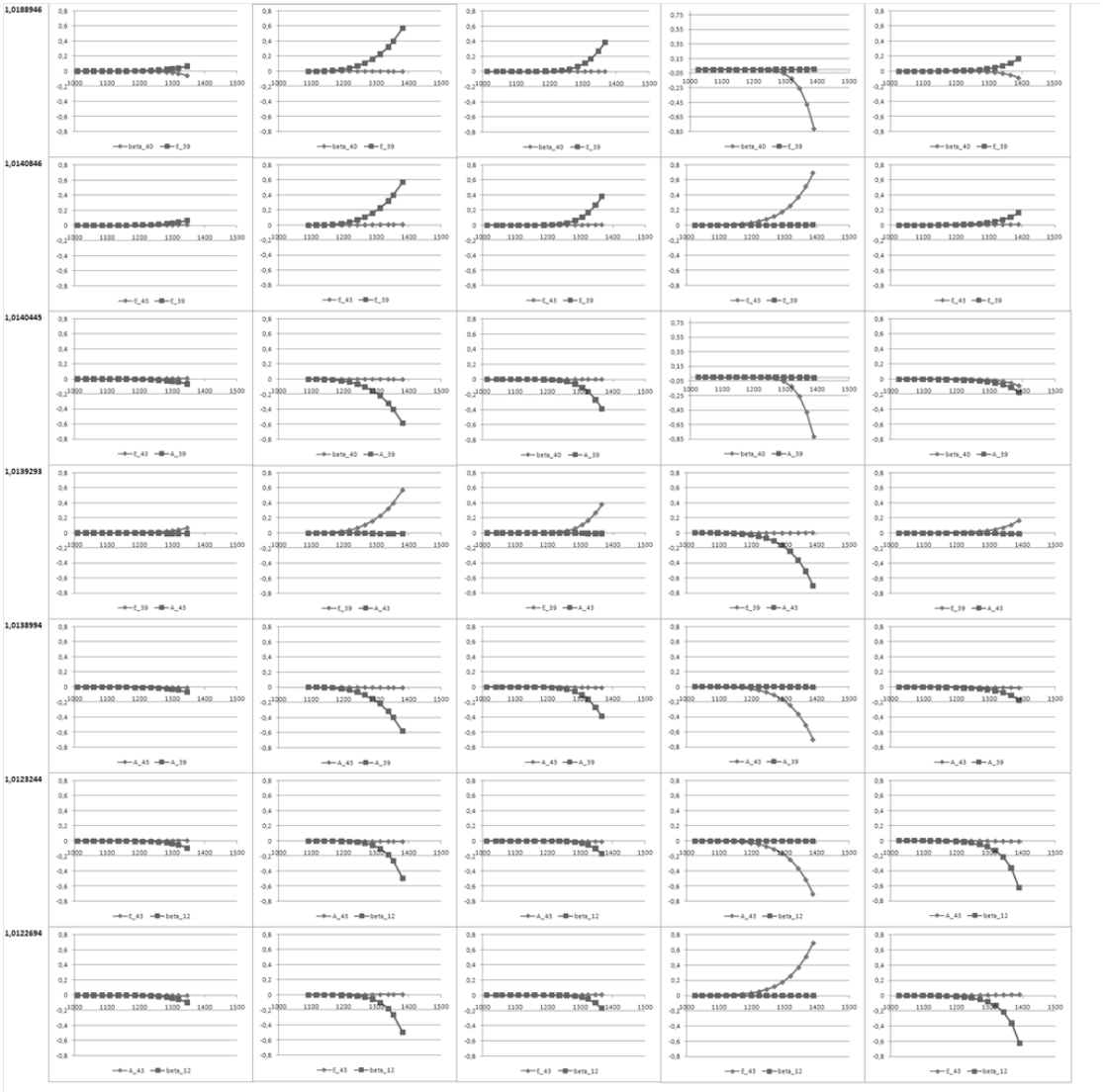


Figure A4.2 *Comparison of sensitivity functions of parameter pairs (here for all parameter pairs that result to be identifiable in the later quantitative analysis, see Chapter 5.1, Phase II, Step 3).*

A5. Discrete diameter fraction measured for limonene aerosol

Table A5.1 *Discrete diameter fractions measured for limonene aerosol*

Number i	***Diameter interval***	***Mean diameter Si***
0	0-1.166µm	0.583
1	1.166-1.359 µm	0.68
2	1.359-1.585 µm	1.47
3	1.585-1.848 µm	1.72
4	1.848-2.154 µm	2.00
5	2.154-2.512 µm	2.33
6	2.512-2.929 µm	2.72
7	2.929-3.415 µm	3.17
8	3.415-3.981 µm	3.70
9	3.981-4.642 µm	4.31
10	4.642-5.412 µm	5.03
11	5.412-6.310 µm	5.86
12	6.310-7.356 µm	6.83
13	7.356-8.577 µm	7.97
14	8.577-10.000 µm	9.29
15	10.000-11.659 µm	10.83
16	11.659-13.594 µm	12.63
17	13.594-15.849 µm	14.72
18	15.849-18.479 µm	17.16
19	18.479-21.544 µm	20.01
20	21.544-25.119 µm	23.33
21	25.119-29.287 µm	27.20
22	29.287-34.146 µm	31.72
23	34.146-39.811 µm	36.98
24	39.811-46.416 µm	43.11
25	46.416-54.117 µm	50.27
26	54.117-63.096 µm	58.61
27	63.096-73.564 µm	68.33
28	73.564-85.770 µm	79.67
29	85.770-100.000 µm	92.88
30	100.000-116.592 µm	108.30
31	116.592-135.936 µm	126.26

A6. Pharmacokinetic case study: Experimental data for CyA concentration in organs and blood

Table A6.1 *Experimental data for CyA concentration in organs [µg/mL] (based on Tanaka et al., 2000)*

time [h]	*cQ2*	*cQ3*	*cQ4*	*cQ5*	*cQ6*	*cQ7*	*cQ8*	*cQ9*	*cQ10*	*cQ11*	*cQ12*	*cQ13*
0.066667	77.2251	34.6222	0.6848	2.2111	5.5563	6.4573	23.4888	24.8083	1.2373	3.3055	21.4187	2.1673
0.066667	77.9105	34.6064	0.6502	1.9921	7.4714	6.7478	20.6366	27.6379	0.795	2.8851	23.153	2.2414
0.066667	79.2961	36.5827	0.6924	-0.6364	4.2355	6.3225	22.729	26.8152	1.6778	3.8537	19.7501	2.0545
0.533333	42.1852	30.8142	0.7088	6.8698	6.3904	14.7483	19.1286	20.68	3.7919	3.4887	25.079	3.7453
0.533333	43.4976	31.5855	0.688	5.4413	6.8124	12.9306	20.1451	18.1624	3.1491	3.7643	26.9251	4.7247
0.533333	44.6612	26.4187	0.7322	3.9565	5.9972	9.7695	18.0116	21.5413	4.0397	3.8463	23.0619	6.461
2.033333	34.6686	25.9534	0.5398	6.6742	6.2707	14.572	12.2168	16.7526	2.9244	4.661	19.4685	8.4468
2.033333	32.4649	23.214	0.5721	8.944	6.4596	13.4374	12.7396	14.7383	3.5526	3.7552	17.1718	8.0059
2.033333	32.855	23.4365	0.4364	9.8788	7.1263	13.3396	16.2996	14.5108	3.469	4.2045	20.645	8.9216
8.033333	14.5341	13.6069	0.3254	3.8963	10.2872	7.2336	5.2745	5.7679	1.1561	4.2098	8.6719	6.3455
8.033333	16.6758	13.3791	0.2434	4.0525	10.2912	5.2973	5.3311	5.2226	1.2751	3.9609	8.5846	7.9923
8.033333	13.6731	13.2921	0.4035	4.1592	9.4217	8.4286	4.5101	5.1189	1.6843	2.3127	8.2027	8.0417
24.03333	5.0925	2.5991	0.1407	1.7366	7.8831	1.5632	0.9531	1.8559	0.4973	2.1082	2.5621	4.9999
24.03333	5.0639	4.1034	0.1398	1.3117	7.9815	2.3505	0.9287	2.2097	0.485	1.8907	2.9452	5.3256
24.03333	4.4574	4.7689	0.2214	0.9651	6.2064	1.7854	2.1844	2.0317	0.5216	2.0962	3.2956	8.3407

Table A6.2 *Experimental data for CyA concentration in blood [µg/mL] (based on Tanaka et al., 2000)*

time	*c_blood_tot*	*c_blood_tot*	*c_blood_tot*	*c_blood_tot*	*c_blood_tot*	*c_blood_tot*
0.0166667	33.6826	30.5112	32.5763	/	/	/
0.0333333	20.3762	21.5643	19.7794	/	/	/
0.05	15.0982	13.035	15.287	/	/	/
0.0666667	8.0784	8.0857	7.8467	/	/	/
0.1166667	6.3069	6.535	6.6195	/	/	/
0.2	5.6823	5.6813	5.1346	/	/	/
0.3666667	3.5988	4.2409	4.0169	/	/	/
0.5333333	2.8425	2.8385	3.7555	3.1279	2.8774	4.0462
1.0333333	3.4215	3.4154	2.572	/	/	/
1.5333333	2.5084	2.5605	2.648	/	/	/
2.0333333	1.8606	1.8046	2.0141	2.1368	1.6011	1.54
4.0333333	1.4733	1.4971	1.2874	/	/	/
6.0333333	1.0479	1.0605	1.7478	/	/	/
8.0333333	1.3946	1.3926	1.3629	1.2673	1.1807	1.2218
12.033333	1.0139	1.1236	0.8724	/	/	/
16.5039	0.7186	0.7427	0.738	/	/	/
24.033333	0.3387	0.3491	0.3754	/	/	/

A7. Model discrimination method (Stewart at al., 1996; 1998)

The model candidates M_j are discriminated based on their posterior probability $p^*(M_j|\boldsymbol{Y},\boldsymbol{\Sigma})$ to be the mathematical expectation of the experimental data $\boldsymbol{Y}$ with the covariance matrix $\boldsymbol{\Sigma}$. In general the posterior probability of a model candidate M_j can be calculated applying Bayes' theorem:

$$p^*(M_j|\boldsymbol{Y},\boldsymbol{\Sigma}) = p_0(M_j)p(\boldsymbol{Y}|M_j,\boldsymbol{\Sigma})/C \quad \text{(A7.1)}$$

Here $p_0(M_j)$ is the apriori probability of the model j. C is an proportionality constant.

Since the models contain unknown model parameter Equation A7.1 results in:

$$p^*(M_j|\boldsymbol{Y},\boldsymbol{\Sigma}) = p_0(M_j)p_0(\boldsymbol{\theta_j}|Mj)\int p(\boldsymbol{Y}|\boldsymbol{\theta_j},M_j,\boldsymbol{\Sigma})d\boldsymbol{\theta_j}^{\boldsymbol{e}}/C \quad \text{(A7.2)}$$

Here $p_0(\boldsymbol{\theta_j}|Mj)$ is the apriori probability of the parameter vector $\boldsymbol{\theta_j}$ of model M_j and $p(\boldsymbol{Y}|\boldsymbol{\theta_j},M_j,\boldsymbol{\Sigma})$ is the probability of $\boldsymbol{Y}$ conditional to $\boldsymbol{\theta_j},M_j,\boldsymbol{\Sigma}$ which is identical to the likelihood function $l(\boldsymbol{\theta_j}|\boldsymbol{Y},M_j,\boldsymbol{\Sigma})$ of the parameter vector $\boldsymbol{\theta_j}$ conditional to the experimental data and the model. The integration (summation over all allowed parameter values $\boldsymbol{\theta_j}^{\boldsymbol{e}}$) is needed in order to eliminate the effect of specific parameter values from the posterior probability of the model candidate M_j.

Stewart at al. (1996, 1998) have made some simplifications and assumptions to derive an expression for the posterior probability $p^*(M_j|\boldsymbol{Y},\boldsymbol{\Sigma})$ in Equation A7.2. For example, the error distribution of the experimental data points is assumed to be normal which results in the following likelihood function:

$$p(\boldsymbol{Y}|\boldsymbol{\theta_j},M_j,\boldsymbol{\Sigma}) = l(\boldsymbol{\theta_j}|\boldsymbol{Y},M_j,\boldsymbol{\Sigma}) = |2\pi\boldsymbol{\Sigma}|^{-n/2}exp\left\{-\frac{1}{2}\sum_{u=1}^{n}\sum_{i=1}^{m}\sum_{k=1}^{m}\left[Y_{iu}-F_{iu}(\boldsymbol{\theta_j})\right]\sigma^{ik}\left[Y_{ku}-F_{ku}(\boldsymbol{\theta_j})\right]\right\} \quad \text{(A7.3)}$$

Here, σ^{ik} is the ik-element of $\boldsymbol{\Sigma}^{-1}$. Further simplifications like for example the linearization of the integrant in Equation A7.2 around the least square estimates of the parameters $\boldsymbol{\theta_j}$ results in simplified expressions for $p^*(M_j|\boldsymbol{Y},\boldsymbol{\Sigma})$ for four different cases:

Case 1-a) Single measured variable, known covariance (of measurement errors)

$$p^*(M_j|\boldsymbol{Y},\sigma) \propto p_0(M_j) * 2^{-\frac{p_j}{2}} * exp\left(-\frac{\hat{S}_j}{2\sigma^2}\right),\ j=1,\dots,n \quad \text{(A7.4)}$$

With: p_j – number of unknown model parameters of model j

$\hat{S}_j$ – minimum sum of least squares

Case1-b) Single measured variable, unknown covariance

$$p^*\left(M_j|\boldsymbol{Y}, S_e, \nu_e\right) \propto p_0\left(M_j\right) * 2^{-\frac{p_j}{2}} * \hat{S}_j^{-\nu_e/2},\ j = 1, \dots, n \quad \text{(A7.5)}$$

With: p_j – number of unknown model parameters of model j

$\hat{S}_j$ – minimum sum of least squares

ν_e – degree of freedom of covariance estimation

Case 2-a) Multiple measured variables, known covariance matrix

$$p^*\left(M_j|\boldsymbol{Y}, \boldsymbol{\Sigma}\right) \propto p\left(M_j\right) * 2^{-\frac{p_j}{2}} * exp\left(-\frac{\hat{S}_j}{2}\right),\ j = 1, \dots, n \quad \text{(A7.6)}$$

With: $\hat{S}_j$ is the sum of squares of the vector $\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{E_u}(\hat{\theta}_j)$, $\boldsymbol{E}_u$-error vector for datapoint u

Case 2-b) Multiple measured variables, unknown covariance matrix

$$p^*\left(M_j|\boldsymbol{Y}, \boldsymbol{\Sigma}\right) = p_0(M_j) * 2^{-\frac{p_j}{2}} * \left|\hat{v}_j\right|^{-\nu_e/2} \quad \text{(A7.7)}$$

With: $\left|\hat{v}_j\right|$ - minimized value of the determinant of the experimental error moment matrix $\boldsymbol{v}(\boldsymbol{\theta_j})$:

$$v_{ik}(\boldsymbol{\theta_j}) = \sum_{u=1}^{n}\left[Y_{iu} - F_{ji}(\boldsymbol{\xi_u}, \boldsymbol{\theta_j})\right]\left[Y_{uk} - F_{jk}(\boldsymbol{\xi_u}, \boldsymbol{\theta_j})\right] \quad \text{(A7.8)}$$

A8. Pharmacokinetic case study: Model equations and numerical model analysis for multi-scale scenario 3.

A8.1 Model equations

The model equations for scenario 3 are given below:

$$finj = \begin{cases} if\ 0 \le t \le 2\ min\text{:}\ \frac{m_{inj} \cdot m_{rat}}{t_{inj} \cdot V_b(47) \cdot fup} \\ \quad if\ t > 2\ min\text{:}\ 0 \end{cases} \quad \text{(A8.1)}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\left(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\right) - kph \cdot c_p(i) + khp \cdot c_h(i) \cdot \frac{fuh}{fup} - kpa \cdot c_p(i) + kap \cdot c_a(i)\frac{fua}{fup}, \quad i = 1, 14-46, 48-57 \quad \text{(A8.2- A8.45)}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\left(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\right) - kph \cdot c_p(i) + khp \cdot \frac{fuh}{fup} \cdot c_h(i) - kpa \cdot c_p(i) + kap \cdot \frac{fua}{fup} \cdot c_a(i) - kpQ(i) \cdot \frac{1}{V_b(i) \cdot fup} \cdot (c_p(i) - c_Q(i)), \quad i = 2-13 \quad \text{(A8.46- A8.57)}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\left(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\right) - kph \cdot c_p(i) + khp \cdot c_h(i) \cdot \frac{fuh}{fup} - kpa \cdot c_p(i) + kap \cdot c_a(i)\frac{fua}{fup} + finj, \quad i = 47 \quad \text{(A8.58)}$$

$$\frac{dc_a(i)}{dt} = \frac{fua}{V_b(i)}\left(fl_{in}(i) \cdot c_{a,in}(i) - fl_{out}(i) \cdot c_a(i)\right) + kpa \cdot \frac{fup}{fua} \cdot c_p(i) - kap \cdot c_a(i), i = 1-57 \quad \text{(A8.59- A8.115)}$$

$$\frac{dc_h(i)}{dt} = \frac{fuh}{V_b(i)}\left(fl_{in}(i) \cdot c_{h,in}(i) - fl_{out}(i) \cdot c_h(i)\right) + kph \cdot \frac{fup}{fuh} \cdot c_p(i) - khp \cdot c_h(i), i = 1-57 \quad \text{(A8.116- A8.172)}$$

$$\frac{dc_Q(j)}{dt} = kpQ(j) \cdot \frac{1}{V_Q(j)} \cdot (c_p(j) - c_Q(j)) - ke(j-1) \cdot c_Q(j), \quad j = 2, 3 \quad \text{(A8.173, A8.174)}$$

$$\frac{dc_Q(j)}{dt} = kpQ(j) \cdot \frac{1}{V_Q(j)} \cdot (c_p(j) - c_Q(j)), \quad j = 4-13 \quad \text{(A8.175- A8.184)}$$

$$\frac{dc_m(l)}{dt} = ke(l) \cdot c_Q(l+1), \quad l = 1, 2 \quad \text{(A8.185, A8.186)}$$

$$V_{blood} = \sum_{i=1}^{57} V_b(i) \quad \text{(A8.187)}$$

$$c_{plasmaU} = \frac{fup}{(fup+fua)V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_p(i) \quad \text{(A8.188)}$$

$$c_{hct} = \frac{1}{V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_h(i) \quad \text{(A8.189)}$$

$$c_{plasmaB} = \frac{fua}{(fup+fua)V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_a(i) \quad \text{(A8.190)}$$

$$c_{plasmaT} = c_{plasmaB} + c_{plasmaU} \quad \text{(A8.191)}$$

$$c_{blood\_tot} = \left(fuh \cdot c_{hct} + (fua + fup) \cdot c_{plasmaT}\right) \quad \text{(A8.192)}$$

$$m_{inVivo} = fup \cdot c_{plasmaU} \cdot V_{blood} + fuh \cdot c_{hct} \cdot V_{blood} + fua \cdot c_{plasmaB} \cdot V_{blood} + \sum_{j=2}^{13} V_Q(j) \cdot c_Q(j) \quad \text{(A8.193)}$$

The model equations can be introduced to the modelling tool ICAS-MoT in a text format and are subsequently translated and analysed by the tool. In this case the model is already available in the model library.

A8.2 Model analysis (numerical)

The results of the numerical model analysis for the high-detail-scale model are provided here.

1) Classification of equations

8 explicit algebraic equations
185 ordinary differential equations

2) Pre-classification of variables in dependent, independent and algebraic variables

Table A8.1 *Pre-classification of model variables in dependent, independent and algebraic variables*

type	***number***	***variables***
dependent	185	$c_p(i)$, $c_a(i)$, $c_h(i)$, i=1-57 $c_Q(j)$, j=2-13 $c_m(l)$, l=1,2
independent	1	t
general variables	170	$finj, m_{inj}, m_{rat}, t_{inj}, fup, fua, fuh, kph, khp, kpa, kap,$ $V_{blood}, c_{plasmaU}, c_{hct}, c_{plasmaB}, c_{plasmaT}, c_{blood\_tot}, m_{inVivo}$ $V_b(i)$, $fl(i)$, i=1-57 $kpQ(j)$, $V_Q(j)$, j=2-13 $ke(l)$, l=1,2

3) Degree of freedom analysis

a) AE part:

$DOF_{AE} = algebraic\ variables\ in\ AEs - number\ of\ AEs$ =83-8=75

Consequently, 75 algebraic variables that appear in the AEs need to be specified as either known or parameter in order to satisfy the degree of freedom. In this case all 75 variables are specified as known: $m_{inj}, m_{rat}, t_{inj}, fup, fua, fuh, V_b(i), V_Q(j)$.

b) ODE part:

$DOF_{ODE} = algebraic\ variables\ that\ do\ not\ appear\ in\ AEs = 75$

These 75 variables need to be specified as either known or parameter to satisfy the degree of freedom:

- 18 variables are specified as parameters: kph, khp, kpa, kap, $kpQ(j)$, $ke(l)$
- 57 variables are specified as known: $fl(i)$ (these are the steady state blood flows which are known from the solution of the lower degree of detail model).

4) Provide variable values

Table A8.3 gives an overview from which sources the known model variables are obtained.

Table A8.3 Values for known variables and their s*ources (Scenario 3)*

Known variable	***value***	***source***	***comment***
m_{inj}	6 mg/kg rat	Tanaka et al., 2000	Corresponds to conditions of experimental data
m_{rat}	251.36 g	Tanaka et al., 2000	Corresponds to conditions of experimental data, less than 277 g because not all organs considered in network
t_{inj}	2 min	Tanaka et al., 2000	Corresponds to conditions of experimental data
$V_b(i)$	See mot-file	Mosat et al., 2011	
$V_Q(j)$	See mot-file	Kawai et al., 1998	
fup	0.37	Mosat et al., 2011	Volume fraction of compartment
fua	0.18	Mosat et al., 2011	Volume fraction of compartment
fuh	0.45	Mosat et al., 2011	Volume fraction of compartment
$fl(i)$	See mot-file	Macro scale model	

The initial conditions for the dependent variables are given by:

$c_p(i)(t=0)=0\,,\quad i=1-57$

$c_a(i)(t=0)=0\,,\quad i=1-57$

$c_h(i)(t=0)=0\,,\quad i=1-57$

$c_Q(j)(t=0)=0\,,\quad j=2-13$

$c_m(l)(t=0)=0\,,\quad l=1-2$

5) Incidence matrix

Table A8.2 *Incidence matrix for high-detail-scale model (Scenario 3)*

eq	expl var	$finj$	$c_p(i)$	$c_a(i)$	$c_h(i)$	$c_Q(i)$	$c_m(i)$	V_{blood}	$c_{plasmaU}$	c_{hct}	$c_{plasmaB}$	$c_{plasmaT}$	c_{blood_tot}	m_{inVivo}
1	$finj$	*												
2-58	$c_p(i)$	*	*	*	*									
59-115	$c_a(i)$		*	*										
116-172	$c_h(i)$		*		*									
173-184	$c_Q(i)$		*			*								
185,186	$c_m(i)$					*	*							
187	V_{blood}							*						
188	$c_{plasmaU}$		*					*	*					
189	c_{hct}				*			*		*				
190	$c_{plasmaB}$			*				*			*			
191	$c_{plasmaT}$								*		*	*		
192	c_{blood_tot}									*		*	*	
193	m_{inVivo}							*	*	*	*			*

A9. Pharmacokinetic case study: Model equations and numerical model analysis for multi-scale scenario 4.

A9.1 Model equations

The model equations for scenario 4 are given below:

$$finj = \begin{cases} if\ 0 \le t \le 2\ min\text{:}\ \frac{m_{inj} \cdot m_{rat}}{t_{inj} \cdot V_b(47) \cdot fup} \\ \quad if\ t > 2\ min\text{:}\ 0 \end{cases} \quad \text{(A9.1)}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\Big(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\Big) - kph \cdot c_p(i) + khp \cdot c_h(i) \cdot \frac{fuh}{fup} - kpa \cdot c_p(i) + kap \cdot c_a(i)\frac{fua}{fup}, \quad i = 1, 14-46, 48-57 \quad \text{(A9.2- A9.45)}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\Big(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\Big) - kph \cdot c_p(i) + khp \cdot \frac{fuh}{fup} \cdot c_h(i) - kpa \cdot c_p(i) + kap \cdot \frac{fua}{fup} \cdot c_a(i) - kpQ(i) \cdot \frac{1}{V_b(i) \cdot fup} \cdot (c_p(i) - c_{IF}(i)) + kQp(i) \cdot \frac{1}{V_b(i) \cdot fup} \cdot c_Q(i), \quad i = 2-13 \quad \text{(A9.46- A9.57)}$$

$$\frac{dc_p(i)}{dt} = \frac{fup}{V_b(i)}\Big(fl_{in}(i) \cdot c_{p,in}(i) - fl_{out}(i) \cdot c_p(i)\Big) - kph \cdot c_p(i) + khp \cdot c_h(i) \cdot \frac{fuh}{fup} - kpa \cdot c_p(i) + kap \cdot c_a(i)\frac{fua}{fup} + finj, \quad i = 47 \quad \text{(A9.58)}$$

$$\frac{dc_a(i)}{dt} = \frac{fua}{V_b(i)}\Big(fl_{in}(i) \cdot c_{a,in}(i) - fl_{out}(i) \cdot c_a(i)\Big) + kpa \cdot \frac{fup}{fua} \cdot c_p(i) - kap \cdot c_a(i), i = 1-57 \quad \text{(A9.59- A9.115)}$$

$$\frac{dc_h(i)}{dt} = \frac{fuh}{V_b(i)}\Big(fl_{in}(i) \cdot c_{h,in}(i) - fl_{out}(i) \cdot c_h(i)\Big) + kph \cdot \frac{fup}{fuh} \cdot c_p(i) - khp \cdot c_h(i), i = 1-57 \quad \text{(A9.116-A9.172)}$$

$$\frac{dc_{IF}(j)}{dt} = kpQ(j) \cdot \frac{1}{V_Q(j) \cdot fIF} \cdot \Big(c_p(j) - c_{IF}(j)\Big) - kTp(j) \cdot c_{IF}(j) + kpT(j) \cdot c_{TC}(j), i = 2-13 \quad \text{(A9.173-A9.184)}$$

$$\frac{dc_{TC}(j)}{dt} = kTp(j) \cdot c_{IF}(j) - kpT(j) \cdot c_{TC}(j) - ke(j-1) \cdot c_{TC}(j), \quad j = 2, 3 \quad \text{(A9.185- A9.186)}$$

$$\frac{dc_{TC}(j)}{dt} = kTp(j) \cdot c_{IF}(j) - kpT(j) \cdot c_{TC}(j), \quad j = 4-13 \quad \text{(A9.187 – A9.196)}$$

$$\frac{dc_m(l)}{dt} = ke(l) \cdot c_Q(l+1), \quad l = 1, 2 \quad \text{(A9.197, A9.198)}$$

$$V_{blood} = \sum_{i=1}^{57} V_b(i) \quad \text{(A9.199)}$$

$$c_{plasmaU} = \frac{fup}{(fup+fua)V_{blood}} \sum_{k=1}^{57} V_b(i) \cdot c_p(i) \quad \text{(A9.200)}$$

$$c_{hct} = \frac{1}{V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_h(i) \quad \text{(A9.201)}$$

$$c_{plasmaB} = \frac{fua}{(fup+fua)V_{blood}} \sum_{i=1}^{57} V_b(i) \cdot c_a(i) \quad \text{(A9.202)}$$

$$c_{plasmaT} = c_{plasmaB} + c_{plasmaU} \quad \text{(A9.203)}$$

$$c_{blood\_tot} = \big(fuh \cdot c_{hct} + (fua + fup) \cdot c_{plasmaT}\big) \quad \text{(A9.204)}$$

$m_{inVivo} = fup \cdot c_{plasmaU} \cdot V_{blood} + fuh \cdot c_{hct} \cdot V_{blood} + fua \cdot c_{plasmaB} \cdot V_{blood} + \sum_{j=2}^{13} V_Q(j) \cdot c_Q(j)$ (A9.205)

The model for scenario 4 is available in the model library.

A9.2 Model analysis (numerical)

The results of the numerical model analysis for the high-detail-scale model are provided here.

1) Classification of equations

8 explicit algebraic equations

197 ordinary differential equations

2) Pre-classification of variables in dependent, independent and general variables

Table A9.1 *Pre-classification of model variables in dependent, independent and algebraic variables*

type	***number***	***variables***
dependent	197	$c_p(i)$, $c_a(i)$, $c_h(i)$, i=1-57 $c_{IF}(j)$, $c_{TC}(j)$, j=2-13 $c_m(l)$, l=1,2
independent	1	t
general variables	183	$finj, m_{inj}, m_{rat}, t_{inj}, fup, fua, fuh, kph, khp, kpa, kap, fIF$ $V_{blood}, c_{plasmaU}, c_{hct}, c_{plasmaB}, c_{plasmaT}, c_{blood\_tot}, m_{inVivo}$ $V_b(i)$, $fl(i)$, i=1-57 $kpQ(j), kpT(j), kTp(j), V_Q(j)$, j=2-13 $ke(l)$, l=1,2

3) Degree of freedom analysis

a) AE part:

$DOF_{AE} = algebraic\ variables\ in\ AEs - number\ of\ AEs$ =83-8=75

Consequently, 75 algebraic variables that appear in the AEs need to be specified as either known or parameter in order to satisfy the degree of freedom. In this case all 75 variables are specified as known: $m_{inj}, m_{rat}, t_{inj}, fup, fua, fuh, V_b(i), V_Q(j)$.

b) ODE part:

$DOF_{ODE} = algebraic\ variables\ that\ do\ not\ appear\ in\ AEs = 90$

These 90 variables need to be specified as either known or parameter to satisfy the degree of freedom:

- 43 variables are specified as parameters: kph, khp, kpa, kap, $kpQ(j)$, $kpT(j)$, $kTp(j)$, fIF, $ke(l)$
- 57 variables are specified as known: $fl(i)$ (these are the steady state blood flows which are known from the solution of the lower degree of detail model).

4) Provide variable values

Table A9.3 gives an overview from which sources the known model variables are obtained.

Table A9.3 Values for known variables and their s*ources (Scenario 3)*

Known variable	***value***	***source***	***comment***
$\boldsymbol{m_{inj}}$	6 mg/kg rat	Tanaka et al., 2000	Corresponds to conditions of experimental data
$\boldsymbol{m_{rat}}$	251.36 g	Tanaka et al., 2000	Corresponds to conditions of experimental data, less than 277 g because not all organs considered in network
$\boldsymbol{t_{inj}}$	2 min	Tanaka et al., 2000	Corresponds to conditions of experimental data
$\boldsymbol{V_b(i)}$	See mot-file	Mosat et al., 2011	
$\boldsymbol{V_Q(j)}$	See mot-file	Kawai et al., 1998	
$\boldsymbol{fup}$	0.37	Mosat et al., 2011	Volume fraction of compartment
$\boldsymbol{fua}$	0.18	Mosat et al., 2011	Volume fraction of compartment
$\boldsymbol{fuh}$	0.45	Mosat et al., 2011	Volume fraction of compartment
$\boldsymbol{fl(i)}$	See mot-file	Macro scale model	

The initial conditions for the independent variables are given below:

$c_p(i)(t=0)=0\,,\quad i=1-57$

$c_a(i)(t=0)=0\,,\quad i=1-57$

$c_h(i)(t=0)=0\,,\quad i=1-57$

$c_{lF}(j)(t=0)=0\,,\quad j=2-13$

$c_{TC}(j)(t=0)=0\,,\quad j=2-13$

$c_m(i)(t=0)=0\,,\quad l=1-2$

5) Incidence matrix

Table A9.2 *Incidence matrix for high-detail-scale model (Scenario 3)*

eq	expl var	$finj$	$c_p(i)$	$c_a(i)$	$c_h(i)$	$c_{IF}(j)$	$c_{TC}(j)$	$c_m(i)$	V_{blood}	$c_{plasmaU}$	c_{hct}	$c_{plasmaB}$	$c_{plasmaT}$	c_{blood_tot}	m_{inVivo}
1	$finj$	*													
2-58	$c_p(i)$	*	*	*	*										
59-115	$c_a(i)$		*	*											
116-172	$c_h(i)$		*		*										
173-184	$c_{IF}(j)$		*			*	*								
185-196	$c_{TC}(j)$					*	*								
197,198	$c_m(i)$						*	*							
199	V_{blood}								*						
200	$c_{plasmaU}$		*						*	*					
201	c_{hct}				*				*		*				
202	$c_{plasmaB}$			*					*			*			
203	$c_{plasmaT}$									*		*	*		
204	c_{blood_tot}										*		*	*	
205	m_{inVivo}								*	*	*	*			*

Nomenclature

A	-	dimensionless coefficient [-]
$A[k]$	-	1st Arrhenius parameter for rate of forward reaction of reaction k [mol/(m^3 s K$^{\beta}$)]
$a1[j]$, …, $a7[j]$	-	parameters for Nasa polynomials of component j [J/(mol K^x)]
aak	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [-]
$Alowk$	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [mol/(m^3 s K$^{\beta}$)]
a_v [1/m]	-	interfacial area between the catalyst and the bulk fluid per unit volume
b	-	parameter of Langmuir isotherm [m^3/kg]
c	-	concentration of adsorbed protein in particle pores [kg/m^3]
C	-	dimensionless coefficient [-]; or proportionality constant [-]
$c_a(i)$	-	concentration of drug in bound plasma compartment cell i [kg/m^3]
$c_{a,in}$	-	concentration of drug in stream entering bound plasma compartment [kg/m^3] cell i
C_a	-	concentration of compound a in droplet [-]
c_b	-	concentration of drug in blood [kg/m^3]
$c_{p,in}$	-	concentration of drug in stream entering blood compartment cell I [kg/m^3]
c_{blood_tot}	-	mean drug concentration in overall blood [kg/m^3]
c_g	-	gas heat capacity [J/(kgK)]
$c_h(i)$	-	concentration of drug in hematocrit compartment cell i [kg/m^3]
c_{hct}	-	mean drug concentration in overall plasma hematocrit [kg/m^3]
$c_{h,in}$	-	concentration of drug in stream entering hematocrit compartment [kg/m^3] cell i
$c_{IF}(i)$	-	mean drug concentration in intersticial fluid (IF) of organ i [kg/m^3]
ck	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [-]
$c_m(i)$	-	concentration of metabolite created by metabolic reaction i [kg/m^3]
$c_p(i)$	-	concentration of drug in unbound plasma compartment cell i [kg/m^3]

$c_{p,in}$	-	concentration of drug in stream entering unbound plasma compartment cell I [kg/m^3]
c_P^{bulk}	-	protein bulk concentration [kg/m^3]
$c_P^{bulk,start}$	-	initial value for protein bulk concentration [kg/m^3]
$c_{plasmaB}$	-	mean drug concentration in overall plasma bound [kg/m^3]
$c_{plasmaT}$	-	mean drug concentration in overall plasma [kg/m^3]
$c_{plasmaU}$	-	mean drug concentration in overall plasma unbound [kg/m^3]
$cp_{liq,i}$ [J/kgK]	-	heat capacity of liquid within droplet of discrete droplet size fraction i
$c_{p\_tot}(i)$	-	concentration of drug in total plasma compartment [kg/m^3] cell i
$cp_{vap,i}$ [J/kgK]	-	heat capacity of vapor from droplet of discrete droplet size fraction i
$c_Q(i)$	-	mean drug concentration in organ i [kg/m^3]
c_s	-	heat capacity of the particles [J/(kgK)]
$c_{TC}(i)$	-	mean drug concentration in tissue cells of organ i [kg/m^3]
$c_{\%,exp}(k)$	-	experimental percentage of droplet size fraction for data point k [%]
$c_{\%,sim}(k)$	-	simulated percentage of droplet size fraction for data point k [%]
d, d_i	-	diameter of droplets, diameter of discrete droplet size fraction i [m]
$D_a D_p$	-	ratio of apparent and pore diffusivity [-]
D_{air}	-	diffusion coefficient in surrounding gas phase [m/s]
D_{disp}	-	dispersion coefficient for droplets [m^2/s]
$E[k]$ [J/mol]	-	3rd Arrhenius parameter for rate of forward reaction of reaction k
$Elowk$	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [J/mol]
F	-	dimensionless coefficient [-]
$F[j], F_j$	-	concentration of component j [mol/s]
$Fcentk$	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [-]
F_i	-	evaporating mass stream from droplets of discrete droplet size fraction i [kg/s]
fIF	-	volume fraction of interstitial fluid of organ
$finj$	-	rate of drug injection [kg/s]
$F_{iu}(\boldsymbol{\theta_j})$	-	model prediction corresponding to variable i of experiment u and parameter vector $\boldsymbol{\theta_j}$

Fk	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [-]
$fl_{in}(i)$		total flow in blood vessel compartment i [m^3/s]
$fl_{out}(i)$		total flow out of blood vessel compartment i [m^3/s]
F_M	-	concentration of 3rd body (for third body enhancement of reactions) [mol/s]
F_{Mk}	-	concentration of 3rd body (specifically for third body enhancement of reaction k) [mol/s]
$F_{N_2O}(i)$	-	simulated concentration of N_2O for conditions of data-point i [mol/s]
$F_{N_2O}^{exp}(i)$	-	measured concentration of N_2O for data-point I [mol/s]
F_{NO}	-	concentration of NO [mol/mol]
fua	-	volume fraction of plasma bound compartment [-]
fuh	-	volume fraction of hematocrit compartment [-]
fup	-	volume fraction of plasma unbound compartment [-]
fuP	-	ratio between the plU and the total plasma concentration [-]
G, G_i	-	growth rate diameters of droplets, growth rate diameter of discrete droplet size fraction i [m/s]
$H^0[j]$, H_j^0	-	standard enthalpy of component j [J/mol]
h_g	-	heat transfer coefficient between the gas and the catalyst particles [$J/(m^2sK$]
H_g	-	dimensionless coefficient [-]
$\mathrm{H_g^*}$	-	dimensionless coefficient [-]
$HR[k]$, HR_k	-	reaction enthalpy of reaction k [J/mol]
H_T	-	dimensionless coefficient [-]
$\mathrm{H_T^*}$	-	dimensionless coefficient [-]
h_w	-	heat transfer coefficient between the reactor wall and the gas [$J/(m^2sK$]
H_w	-	dimensionless coefficient [-]
IS	-	ionic strength [mol/m^3]
k	-	rate of the catalysed reaction [-]
$K[k]$, K_k	-	equilibrium constant of reaction k [-]
K_{air}	-	heat conduction of air [J/(m s K)]
kap	-	kinetic rate constant of drug transfer from plasma bound to plasma unbound[1/s]
$kb[k]$, kb_k	-	rate constant of backward reaction of reaction k [$m^x/(s\ mol^y)$]
K_D	-	Dissociation constant of drug in blood [kg/m^3]
$ke(i)$	-	kinetic rate constant of metabolic reaction i [1/s]

$kf[k]$, kf_k	-	rate constant of forward reaction of reaction k [$m^x/(s\ mol^y)$]
k_g	-	mass transfer coefficient between the bulk fluid and the particles [$mol/(s\ m^2 Pa)$]
khp	-	kinetic rate constant of drug transfer from hermatocrit to plasma [1/s]
$kinfk$	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [$m^x/(s\ mol^y)$]
$klowk$	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [$m^x/(s\ mol^y)$]
kpa	-	kinetic rate constant of drug transfer from plasma unbound to plasma bound [1/s]
kph	-	kinetic rate constant of drug transfer from plasma to hermatocrit [1/s]
$kpQ(i)$	-	kinetic rate constant of drug transfer from plasma unbound to organ i [m^3/s]
$kpT(i)$	-	kinetic rate constant of drug transfer from cell tissue to interstitial fluid of organ i [1/s]
$kQp(i)$	-	kinetic rate constant of drug transfer from organ i to plasma unbound [m^3/s]
$kTp(i)$	-	kinetic rate constant of drug transfer from interstitial fluid to cell tissue of organ i [1/s]
$k_{retention}$	-	chromatographic retention factor [-]
$l(\boldsymbol{\theta_j}\|\boldsymbol{Y}, M_j, \boldsymbol{\Sigma})$	-	likelihood function of the parameter vector $\boldsymbol{\theta_j}$ conditional to the experimental data $\boldsymbol{Y}$, the model M_j and the covariance matrix $\boldsymbol{\Sigma}$
L_i	-	latent heat of droplet with diameter d_i [J/kg]
m_{ads}	-	mass of protein that has been so far stored on the surface and in the particle pores of all particles in the vessel [kg]
m_{ads}^P	-	mass of protein that has been so far stored on the surface and in the particle pores of one particle in the vessel [kg]
m_{inj}	-	mass of injected drug [kg]
m_{inVivo}	-	total mass of drug inside body[kg]
m_{rat}	-	mass of rat [kg]
MW	-	molecular weight [kg/mol]
N	-	number of total data points [-]
N, N_i	-	droplet concentration, concentration of discrete droplet size fraction i [$1/m^3$]
NC	-	number of components in system [-]
NDAT, N_{dat}	-	number of data-points
$Ndis$	-	number of discrete diamters in the droplet size distribution [-]

Nk	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [-]
$NPAR$	-	number of unknown model parameters [-]
nP_T	-	Binding capacity of drug [Eq/m^3]
NR	-	number of reactions [-]
$NVAR$	-	number of response variables for sensitivity analysis [-]
Obj	-	objective function
$\boldsymbol{p}$	-	vector of model parameters
p	-	partial pressure of reactant [Pa]
P	-	pressure [Pa]
$p(\boldsymbol{Y}\|M_j, \boldsymbol{\Sigma})$	-	probability of data $\boldsymbol{Y}$ to be represented by model candidate $M_j$and the covariance matrix $\boldsymbol{\Sigma}$ [-]
$p(\boldsymbol{Y}\|\boldsymbol{\theta_j}, M_j, \boldsymbol{\Sigma})$	-	probability of $\boldsymbol{Y}$ conditional to $\boldsymbol{\theta_j}, M_j, \boldsymbol{\Sigma}$ [-]
P_{am}	-	vapor pressure of droplet compound far away from droplet [Pa]
p_e	-	partial pressure of reactant at reactor entrance [Pa]
P_i^s	-	vapor pressure over droplet surface of discrete droplet size fraction i [Pa]
$P_i^{s,plane}$	-	vapor pressure over plane surface for temperature of discrete droplet size fraction i [Pa]
p_j	-	number of unknown model parameters of model j
P_j	-	base value for model parameter j during sensitivity analysis
p_p	-	partial pressure of reactant at catalyst pellet [Pa]
$p_0(M_j)$	-	apriori probability of model M_j [-]
$p_0(\boldsymbol{\theta_j}\|Mj)$	-	apriori probability of the parameter vector $\boldsymbol{\theta_j}$ of model M_j [-]
$p^*(M_j\|\boldsymbol{Y}, \boldsymbol{\Sigma})$	-	posterior probability of model candidate $M_j$to be the mathematical expectation of the experimental data $\boldsymbol{Y}$ with the covariance matrix $\boldsymbol{\Sigma}$ [-]
q	-	concentration of adsorbed protein on particle surface [kg/m^3]; or gas mass flow rate [kg/s]
q_0	-	equilibrium concentration of adsorbed protein on particle surface [kg/m^3]
q_{ads}	-	total concentration of uptaken protein [kg/m^3]
q_{mon}	-	parameter of Langmuir isotherm [kg/m^3]
r	-	radius [m]
R, Rm	-	general gas constant [J/mol K]
$r[k], r_k$	-	reaction rate of reaction k [mol/(m^3 s)]
R_A	-	agglomeration rates of droplets [1/(s m^3)]

R_B	-	breakage rates of droplets [1/(s m^3)]
$R_{Kelvin,i}$	-	Kelvin factor of droplet with diameter d_i (considers effect of change of vapor pressure over droplet surface compared to plane surface) [-]
$R_{particle}$	-	particle radius [m]
R_{sep}	-	separation factor [-]
$S^0[j]$, S_j^0	-	standard entropy of component j [J/(mol K)]
S_i	-	mean diameter of droplet size fraction i [m]
$\hat{S}_j$	-	minimum sum of least squares
$SR[k]$, SR_k	-	reaction entropy of reaction k [J/(mol K)]
$Sa(i,j)$	-	absolute sensitivity of response variable i with respect to parameter j
$S_{nd}(i,j)$	-	non-dimensional sensitivity of response variable i with respect to parameter j [-]
$S_{nd,K}$	-	matrix of non-dimensional sensitivities for parameter subset K [-]
S_{norm}	-	normalized sensitivity matrix [-]
$S_{norm,ij}$	-	normalized sensitivity of response variable i with respect to parameter j [-]
$S_{norm,K}$	-	normalized sensitivity matrix of parameter subset K [-]
s_p	-	particle surface [m^2]
t	-	independent variable time [s]
T	-	temperature [K]
T_{am}	-	ambient temperature [K]
T_c	-	critical temperature [K]
T_e	-	temperature of stream at reactor entrance[K]
t_{inj}	-	injection time [s]
T_p	-	temperature of catalyst pellets [K]
T_p	-	temperature of wall [K]
$T_{s,i}$	-	temperature at droplet surface of discrete droplet size fraction i [K]
$\boldsymbol{u}$	-	vector formed by all remaining independent variables considered for the system
$\underline{u}$	-	droplet velocity vector [m/s]
V	-	volume [m^3]
V_b	-	volume of blood vessel [m^3]
V_{blood}	-	total blood volume in body [m^3]
V_{bulk}	-	bulk volume [m^3]
V_{cell}	-	volume of discretization cell [m^3]
V_{liq}	-	volume of protein solution [m^3]

Vp	-	volumetric total flow rate [m^3/s]; or catalyst particle volume [m^3]
V_Q	-	volume of organ [m^3]
$V_{suspension}$	-	volume of suspension of stationary phase [m^3]
$V_{particle}$	-	particle volume [m^3]
$\boldsymbol{x}$	-	vector of conditional variables
x	-	x-coordinate of droplets [m]
Xk	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [mol/m^3]
$\boldsymbol{y}$	-	vector of state variables
y	-	y-coordinate of droplets [m]
$\boldsymbol{Y}$	-	experimental data matrix
Y	-	dimensionless concentration of adsorbed protein in particle pores [-], normalized by protein bulk concentration
Y_{Bi}	-	value of response variable i for backward perturbation during sensitivity analysis
Y_{Fi}	-	value of response variable i for forward perturbation during sensitivity analysis
Y_i	-	base value of response variable i during sensitivity analysis
Y_{iu}	-	data-point for measured variable i and experiment u
Y_{iu}	-	data-point for measured variable k and experiment u
$\boldsymbol{z}$	-	represents the vector of known variables
z	-	z-coordinate of droplets [m]
$\underline{z}$	-	vector of droplet position in space [m]

Greek letters

α	-	dimensionless parameter of parallel diffusion model [-]; or Void fraction of particles [-]
β	-	dimensionless parameter of parallel diffusion model [-]
$\beta[k]$, β_k	-	2nd Arrhenius parameter for rate of forward reaction of reaction k [-]
$\beta lowk$	-	parameter for Troe equation (considers pressure dependence of rate constants) of reaction k [-]
γ_K	-	collinearity index of parameter subset K [-]
Γ	-	Stephan-Boltzmann constant [$W/(m^2K^4)$]
δ_j^{msqr}	-	sensitivity measure [-]

ΔH	-	heat of reaction [J/mol]
ΔP_j	-	perturbation of parameter j during sensitivity analysis
ε	-	void fraction of bed [-]
ε_p	-	effective particle porosity [-]
$\boldsymbol{E}_u$	-	error vector for datapoint u
$\boldsymbol{\theta}$	-	vector of constitutive variablesGeben Sie hier eine Formel ein.
κ	-	Debye length [m]
λ	-	eigenvalue [-]
ν_e	-	degree of freedom of covariance estimation
ν_{kj}	-	stoichiometric factor of component j in reaction k [-]
ρ	-	dimensionless radius [-]
ρ_g	-	gas density [kg/m^3]
ρ_i	-	density of droplet with diameter d_i [kg/m^3]
ρ_K	-	determinant measure of parameter subset K [-]
ρ_s	-	particle density [kg/m^3]
σ_i	-	surface tension of droplet with diameter d_i [kg/s^2]
σ^{ik}	-	ik-element of inverse covariance matrix $\boldsymbol{\Sigma}^{-1}$
$\boldsymbol{\Sigma}$	-	covariance matrix
τ_P	-	dimensionless time for protein batch uptake [-]
τ^*	-	dimensionless time for fluidized bed reactor [-]
$\boldsymbol{v}(\boldsymbol{\theta_j})$	-	experimental error moment matrix resulting for parameter vector $\boldsymbol{\theta_j}$ and model j
$v_{ik}(\boldsymbol{\theta_j})$	-	element ik of experimental error moment matrix resulting for parameter vector $\boldsymbol{\theta_j}$ and model j
$\lvert\hat{v}_j\rvert$	-	minimized value of the determinant of the experimental error moment matrix $\boldsymbol{v}(\boldsymbol{\theta_j})$ resulting for parameter vector $\boldsymbol{\theta_j}$ and model j
ϕ_c	-	accessible surface area per unit accessible volume within the chromatographic column [1/m]
ϕ_p	-	accessible surface area per unit particle volume [1/m]

Acronyms

AE	-	Algebraic Equation
BDF	-	Backward Differentiation Formula
BIC	-	Bayesian Information Criterion
CyA	-	Cyclosporin A

discr.	-	discrimination
GAMS	-	General Algebraic Modeling System
gPROMS	-	general Process Modelling System
hct	-	hematocrit (red blood cells)
ICAS	-	Integrated Computer-Aided System
IF	-	Interstitial Fluid
inf.		information
depend.	-	dependent
MAE	-	Mean Absolute Error
DAE	-	Differential Algebraic Equation
MathML	-	Mathematical Markup Language
max.	-	maximal
MLE	-	Maximum Likelihood Estimation
MODEL.LA	-	Modeling Language
MoT	-	Modelling Testbed
ModDev	-	Model Development
NIST	-	National Institute for Standards and Technology
ODE	-	Ordinary Differential Equation
PBPK	-	Physiologically-Based Pharmacokinetic
PDE	-	Partial Differential Equation
RHS	-	Right Hand Side
plB	-	plasma bound
plU	-	plasma unbound
RMSE	-	Root Mean Square Deviation
RPN	-	Reverse Polish Notation
SQP	-	Sequential Quadratic Programming
TC	-	Tissue Cells
wrt.	-	with respect to

References

Aspentech (2003). Aspen Custom Modeler® 12.1 User Guide, Aspen Technology, Inc., Cambridge MA, USA.

Aspentech (2011a). Aspen Custom Modeler, Aspen Technology, Inc. Burlington MA, USA, http://www.aspentech.com/products/aspen-custom-modeler.aspx (accessed 13.11.2011).

Aspentech (2011b). Aspen Plus, Aspen Technology, Inc. Burlington MA, USA, http://www.aspentech.com/core/aspen-plus.aspx (accessed 13.11.2011).

Ausbrooks, R., Buswell, S., Dalmas, S., Devitt, S., Diaz, A., Hunter, R., Smith, B., Soiffer, N., Sutor, R., & Watt, S. (2001). Mathematical Markup Language (MathML) – Version 2.0, available online: http://www.w3.org/TR/MathML2/ (accessed 13.11.2011).

Bahmanyar, H., & Slater, M.J. (1991). Studies of drop break-up in liquid–liquid systems in a rotating disc contactor. Part 1: conditions of no mass transfer. *Chem. Eng. Technol.*, 14, 79-89.

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuiness, D.L., Patel-Schneider, P.F., et al. (2004). OWL web ontology language reference, available online: http://www.w3.org/TR/owl-ref/ (accessed 13.11.2011).

Bogusch, R., Lohmann, B., & Marquardt, W. (2001). Computer-aided process modeling with modkit. *Comput. Chem. Eng.*, 25 (1), 963-995.

Brooke, A., Kendrick, D., Meeraus, A., & Raman, R. (1998). GAMS – A User's Guide. GAMS Development Corp.

Brun, R., Reichert, P., & Künsch, H. R. (2001).Practical identifiability analysis of large environmental simulation models. *Water Resour. Res.*, 37, 1015-1030.

Brun, R., Kühni, M., Siegrist, H., Gujer, W., & Reichert, P. (2002). Practical identifiability of ASM2d parameters – systematic selection and tuning of parameter subsets. *Water Research*, 36, 4113-4127.

Cameron, I., & Gani, R. (2011). *Product and Process Modelling – A Case Study Approach*. 1st ed., Elsevier, Amsterdam.

Carta, G., Ubiera, A. R., & Pabst, T. M. (2005). Protein mass transfer kinetics in ion exchange media: Measurements and interpretations. *Chem. Eng. Technol*, *28*, 1252–1264.

Cauwenberg, V. (1995). Ph.D. Thesis, Katholieke University Leuven, Belgium.

Cháfer, A., Muñoz, R., Burguet, M. C., & Berna, A. (2004). The influence of the temperature on the liquid-liquid equilibria of the mixture water + ethanol + H2O. *Fluid Phase Equilib.*, 224, 251-256.

Costa, C.B.B., da Costa, A.C., & Maciel, R. (2005). Mathematical modeling and optimal control strategy development for an adipic acid crystallization process. *Chem. Eng. Process.*, 44 (7), 737-753.

Coulaloglou, C.A., & Tavlarides, L.L. (1977). Description of interaction processes in agitated liquid-liquid dispersions. *Chem. Eng. Sci.*, 32, 1289-1297.

DePhillips, P., & Lenhoff, A. M. (2000). Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography. *J. Chromatogr., A*, 883, 39–54.

Dziennik, S.R, Belcher, E.B., Barker, G.A., DeBergalis, M.J., Fernandez, S.E., & Lenhoff, A.M. (2003). Nondiffusive mechanisms enhance protein uptake rates in ion exchange particles. *Proc. Na.t Acad. Sci.*, 100, 420-425.

Dziennik, S.R, Belcher, E.B., Barker, G.A., Fernandez, S.E., & Lenhoff, A.M. (2005). Effects of ionic strength on lysozyme uptake rates. *Biotechnol. Bioeng.*, 91, 139-153.

Foss, B., Lohmann, B., & Marquardt, W. (1998). A field study of the industrial modeling process. *J. Process Control*, 5/6(1), 325-338.

Gani, R., Hytoft, G., Jaksland, C., & Jensen, A.K. (1997). An integrated computer aided system for integrated design of chemical processes. *Comput. Chem. Eng.*, 21 (10), 1135-1146.

Glarborg, P., Johnsson, J. E., & Dam-Johansen, K. (1994). Kinetics of Homogeneous Nitrous Oxide Decomposition. *Combust. Flame*, 99, 523-532.

Grossmann, I.E. (2004). Challenges in the New Millennium: Product Discovery and Design, Enterprise and Supply Chain optimization, Global Life Cycle Assessment. *Comput. Chem. Eng.*, 29, 29-39.

Grossmann, I. E., & Westerberg, A.W. (2000). Research challenges in process systems engineering. *AIChE J.*, 46, 1700–1703.

Hall, C., Mosat, A., Lueshen, E. & Linninger, A. (2011) (personal communication).

Heitzig, M., Sin, G., Glarborg, P., & Gani, R. (2010). Computer-aided framework for regression and multi-scale modelling needs in innovative product-process engineering, In: S. Pierucci & G. Buzzi Ferraris (Eds.). *20th European Symposium on Computer Aided Engineering*, Elsevier, Amsterdam, pp. 379-384.

Heitzig, M., Sin, G., Sales-Cruz, M., Glarborg, P., & Gani, R. (2011a). Computer-aided modeling framework for efficient model development, analysis, and identification: combustion and reactor modeling. *Ind. Eng. Chem. Res.*, 50 (9), 5253-5265.

Heitzig, M., Gregson, C., Sin, G., & Gani, R. (2011b). Application of computer-aided multi-scale modelling framework - Aerosol case study, In: E.N. Pistikopoulos, M.C. Georgiadis, & A.C. Kokossis (Eds.). *21st European Symposium on Computer Aided Engineering*, Elsevier, Amsterdam, pp. 16-20.

Heitzig, M., Rong, Y., Gregson, C., Sin, G., & Gani, R. (2012). Systematic multi-scale model development strategy for fragrance spraying process and transport. Accepted by *Chem. Eng. Technol.*

Hulburt, H.M., & Katz, S. (1964). Some Problems in particle technology - A statistical mechanical formulation. *Chem. Eng. Sci.*, 19 (8), 555-574.

Ingram, G.D., Cameron, I.T., & Hangos, K.M. (2004). Classification and analysis of integrating frameworks in multiscale modelling. *Comput. Chem. Eng.*, 59, 2171-2187.

Ingram, G.D. (2005). *Multiscale modelling and analysis of process systems*. *Ph.D. thesis*, Division of Chemical Engineering, University of Queensland.

Jensen, A., & Gani, R. (1996). A computer-aided system for generation of problem specific process models. *Comput. Chem. Eng.* (suppl.), 20, 145-150.

Jensen, A. K. (1998). *Generation of problem Specific Simulation Methods within an Integrated Computer Aided System*. PhD thesis, Department of Chemical Engineering, Technical University of Denmark.

Kawai, R. & Lemaire, M. (1993). Role of blood cell uptake on cyclosporin pharmacokinetics, In: P. Tillement, & H. Eckert (Eds.). *Proceeding of the International Symposium on Blood Binding and Drug Transfer*, EFC Publishing, Paris, pp. 89–108.

Kawai, R., Mathew, D., Tanaka, C., & Rowland, M. (1998). Physiologically based pharmacokinetics of cyclosporine A: Extension to tissue distribution kinetics in rats and scale-up to human. *J. Pharmacol. Exp. Ther.*, 287, 457-468.

Kee, R. J., Rupley, F. M., & Miller, J. A. (1994). The Chemkin Thermodynamic Data Base. *Sandia Report*, SAND87-8215B.

Klatt, K.-U., & Marquardt, W. (2009). Perspectives for process systems engineering – Personal views from academia and industry. *Comput. Chem. Eng.*, 33, 536-550.

Koch, W. *SprayExpo Program Description*. Toxikologie und Experimentelle Medizin, Frauenhofer Institut, Hannover, Germany.

Kukkonen, J., Vesala, T., & Kulmala, M. (1989). The interdependence of evaporation and settling for airborne freely falling droplets. *J. Aerosol Sci.*, 20 (7), 749-763.

Landoldt-Börnstein Database. (2011). Springer Materials.

Lemaire, M. & Tillement, P. (1982). Role of lipoproteins and erythrocytes in the in vitro binding and distribution of cyclosporin A in the blood. *J Pharm. Pharmacol.*,34, 715–718.

Lenhoff, A.M. (2008). Multiscale Modeling of Protein Uptake Patterns in Chromatographic Particles. *Langmuir*, 24, 5991-5995.

Lohmann, B., & Marquardt, W. (1996). On the systematization of the process of model development, *Comput. Chem. Eng.,* 20 (A), S213–S218.

Luss, D., & Amundson, N. R. (1968). Stability of Batch Catalytic Fluidized Beds. *AIChE J.*, 14 (2), 211-221.

Marquardt, W. (1996): Trends in computer-aided process modelling. *Comput. Chem. Eng.*, 20, 591-609.

Modelica Association (2010). Modelica – A Unified Object-Oriented Language for Physical Systems Models. Language Specification Version 3.2, available online: http://www.modelica.org (accessed 13.11.2011).

Morales-Rodríguez, R., & Gani, R. (2007). Computer-Aided Multiscale Modelling for Chemical Process Engineering. *Computer Aided Chemical Engineering*, 24, 207-212.

Morbach, J., Wiesner, A., & Marquardt, W. (2009). OntoCAPE – A (re)usable ontology for computer-aided process engineering. *Comput. Chem. Eng.*, 33, 1546-1556.

Mosat, A., Lueshen, E., Hall, C., & Linninger, A. (2001). First Principles Pharmacokinetic Modeling;Quantitative Study on Cyclosporin A. (unpublished).

Lueshen, E., Hall, C., Mosat, A., & Linninger, A. (2011). Physiologically-based pharmacokinetic modelling: parameter estimation for Cyclosporin A, In: E.N. Pistikopoulos, M.C. Georgiadis, & A.C. Kokossis (Eds.). *21st European Symposium on Computer Aided Engineering*, Elsevier, Amsterdam, pp. 1543-1547.

Nemeth, E., Lakner, R., Hangos, K.M. & Cameron, I.T. (2005). Prediction-based diagnosis and loss prevention using quantitative multi-scale models. In: L. Puigjaner & A. Espuña (Eds.), *15th European symposium on computer-aided process engineering: Vol. CACE 20A-20B,* Elsevier, Amsterdam, pp. 535-540.

National Institute of Standards and Technology (2000). NIST Chemical Kinetics Database, Gaithersburg, MD, USA, available online: http://kinetics.nist.gov/kinetics/index.jsp (21.11.2011).

Pantelides, C. C. (2001). In: R. Gani & S. B. Jorgensen (Eds.). *11th European symposium on computer-aided process engineering*, Elsevier, Amsterdam, pp. 15–26.

Preisig, H.A. (2010). Constructing and maintaining proper process models. *Comput. Chem. Eng.*, 34, 1543-1555.

Process Systems Enterprise (2010a). Model developer guide, gProms release v3.3.0. Process Systems Enterprise Limited.

Process Systems Enterprise (2010b). Model validation guide, gProms release v3.3.0. Process Systems Enterprise Limited.

Process Systems Enterprise (2010c). gProms ModelBuilder guide, gProms release v3.3.0. Process Systems Enterprise Limited.

Process Systems Enterprise (2010d). Optimization guide, gProms release v3.3.0. Process Systems Enterprise Limited.

Process Systems Enterprise (2011). gPROMS, Process Systems Enterprise, http://www.psenterprise.com/gproms/ (accessed 13.11.2011).

Ranz, W. E., & Marshall Jr., W. R. (1952). Evaporation from drops.2.. *Chem. Eng. Prog.*, 48 (4), 173-180.

Renninger, R.G., Hiller, F.C., & Bone, R.C. (1981). The evaporation and growth of droplets having more than one volatile constituent. *J. Aerosol Sci.*, 12 (6), 505-515.

Randolph, A.D., & Larson, M.A. (1988). *Theory of particulate processes*. 2nd ed., Academic Press, San Diego.

Rasmussen, C. L., Hansen, J., Marshall, P., & Glarborg, P. (2008). Experimental Measurements and Kinetic Modeling of $CO/H_2/O_2/NO_X$ Conversion at High Pressure. *Int. J. Chem. Kinet.*, 40, 454-480.

Russel, B. M. R., & Gani, R. (2000). MoT – a modelling test-bed. In: *ICAS Manual (CAPEC Report)*, Technical University of Denmark, Kgs. Lyngby.

Sales-Cruz, M. (2006). *Development of a Computer-Aided Modelling System for Bio and Chemical Process and Product Desgin*. PhD thesis, Department of Chemical Engineering, Technical University of Denmark.

Sales-Cruz, M., & Gani, R. (2003). A modelling tool for different stages of the process life. In: S.P. Asprey, S. Macchietto (Eds.). *Computer-Aided Chemical Engineering, vol 16: Dynamic Model Development*, Elsevier, Amsterdam, pp. 209-249.

Schwarz, G. E. (1978). Estimating the dimension of a model. Ann. Stat., 6 (2), 461–464.

Stephanopoulos, G., Henning, G., & Leone, H. (1990). MODEL.LA – A modeling framework for process engineering - I. The formal framework. *Comput. Chem. Eng.*, 14, 813-846.

Sin, G., & Vanrolleghem, P.A. (2007). Extensions to modeling aerobic carbon degradation using combined respirometric–titrimetric measurements in view of activated sludge model calibration. *Water Research*, 41, 3345-3358.

Sin, G., Eliasson L.A., & Gernaey, K.V. (2009a). Good modeling practice (GMoP) for PAT applications: Propagation of input uncertainty and sensitivity analysis. *Biotechnol. Prog.*, 25, 1043-1053.

Sin, G., Gernaey, K. V., Woodley, J. (2009b). Application of modeling and simulation tools for the evaluation of biocatalytic processes: a future perspective. Biotechnol. Progr., 25 (6), 1529-1538.

Sin, G., Meyer, A. S., Gernaey, K. V. (2010). Assessing reliability of cellulose hydrolysis models to support biofuel process design–identifiability and Uncertainty analysis, *Comput. Chem. Eng.*, 34, 1385–1392.

Simon, M., Schmidt, S.A., & Bart, H.J. (2003). The droplet population balance model - Estimation of breakage and coalescence. *Chem. Eng. Technol.*, 26 (7), 745-750.

Simsci (2011). Pro II, Invensys Operations Management, http://iom.invensys.com/AP/Pages/SimSci-Esscor_ProcessEngSuite_PROII.aspx (accessed 13.11.2011).

Stewart, W.E., Caracotsios, M., & Sørensen, J.P. (1992). Parameter estimation from multiresponse data. *AIChE J.*, 38 (5), 641-650.

Stewart, W.E., Henson, T.L., & Box, G.E.P. (1996). Model discrimination and criticism with single-response data. *AIChE J.*, 42 (11), 3055-3062.

Stewart, W.E., Shon, Y., & Box, G.E.P. (1998). Discrimination and goodness of fit for multiresponse mechanistic models. *AIChE J.*, 44 (6), 1404-1412.

Tanaka, C., Kawai, R., & Rowland, M. (2000). Dose-dependent pharmacokinetics of cyclosporin A in rats: Events in tissues. *Drug Metab. Dispos.*, 28 (5), 582-589.

Troe, J. (1979). Predictive possibilities of unimolecular rate theory. J. *Phys. Chem.*, 83, 114-126.

von Wedel, L., Marquardt, W., & Gani, R. (2002). Modelling frameworks, in: B. Braunschweig, & R. Gani (Eds.). *Software Architectures and Tools for Computer Aided Process Engineering*, Elsevier, Amsterdam, pp. 87-125.

Wilke, C.R., & Lee, C.Y. (1955). Estimation of diffusion coefficients for gases and vapors. *Ind. Eng. Chem. Res.*, 47 (6), 1253-1257.

Yang, A., & Marquardt, W. (2004). An ontology-based approach to conceptual process modeling. In: A. Barbarosa-Póvoa, & H. Matos (Eds.). 14th *European Symposium on Computer Aided Engineering*, Elsevier, Amsterdam, pp. 1159–1164.

Yang, A., & Marquardt, W. (2009). An ontological conceptualization of multiscale models. *Comput. Chem. Eng.*, 33, 822-837.

Zwietering, T.N. (1959). The degree of mixing in continuous flow systems. *Chem. Eng. Sci.*, 11, 1-15.

This PhD-project was carried out at CAPEC, the Computer Aided Product-Process Engineering Center.
CAPEC is committed to research, to work in close collaboration with industry and to participate in educational activities. The research objectives of CAPEC are to develop computer-aided systems for product/process simulation, design, analysis and control/operation for chemical, petrochemical, pharmaceutical and biochemical industries. The dissemination of the research results of CAPEC is carried out in terms of computational tools, technology and application. Under computational tools, CAPEC is involved with mathematical models, numerical solvers, process/operation mathematical models, numerical solvers, process simulators, process/product synthesis/design toolbox, control toolbox, databases and many more. Under technology, CAPEC is involved with development of methodologies for synthesis/design of processes and products, analysis, control and operation of processes, strategies for modelling and simulation, solvent and chemical selection and design, pollution prevention and many more. Under application, CAPEC is actively involved with developing industrial case studies, tutorial case studies for education and training, technology transfer studies together with industrial companies, consulting and many more.

Further information about CAPEC can be found at www.capec.kt.dtu.dk.

Computer Aided Process Engineering Center
Department of Chemical and Biochemical Engineering
Technical University of Denmark
Søltofts Plads, Building 229
DK-2800 Kgs. Lyngby
Denmark

Phone: +45 4525 2800
Fax: +45 4525 4588
Web: www.capec.kt.dtu.dk

ISBN : 978-87-92481-73-3